Air Force Research Laboratory Warfighter Readiness Research Division Participation in the 2008 IITSEC

Compiled by:

Elizabeth P. Casey

NCI Information Systems, Inc
6030 South Kent Street
Mesa AZ 85212-6061

December 2008

Air Force Research Laboratory
Human Effectiveness Directorate
Warfighter Readiness Research Division

NOTICE

This paper is published in the interest of scientific and technical information exchange and its publication does not constitute the Government's approval or disapproval of its idea or findings.

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

Requestors may obtain copies of this report from the Defense Technical Information Center (DTIC) at http://www.dtic.mil.

These papers have been cleared for public release by the Office of Public Affairs, have been presented in an open forum, and published in the proceedings with an unlimited distribution. No further clearance is required.

AFRL-RH-AZ-TP-2008-0002 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

/signed/

HERBERT H. BELL
Technical Advisor

/signed/

DANIEL R. WALKER, Colonel, USAF
Chief, Warfighter Readiness Research Division
Air Force Research Laboratory

REPORT DOCUMENTATION PAGE

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.**

1. REPORT DATE *(DD-MM-YYYY)* 12-15-2008	2. REPORT TYPE Proceedings		3. DATES COVERED *(From - To)* 1 January 2008 – 3 December 2008
4. TITLE AND SUBTITLE Air Force Research Laboratory Warfighter Readiness Research Division Participation in IITSEC 2008			5a. CONTRACT NUMBER F41624-05-D-6502
			5b. GRANT NUMBER
			5c. PROGRAM ELEMENT NUMBER 62205F
6. AUTHOR(S) Compiler: Elizabeth P. Casey			5d. PROJECT NUMBER 1123
			5e. TASK NUMBER AM
			5f. WORK UNIT NUMBER 01
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) NCI Inc 6030 South Kent Street Mesa AZ 85212-6061			8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory Human Effectiveness Directorate Warfighter Readiness Research Division 6030 South Kent Street Mesa AZ 85212-6061			10. SPONSOR/MONITOR'S ACRONYM(S) AFRL; AFRL/RHA
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) AFRL-RH-AZ-TP-2008-0002

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
Attached papers were presented at the 2008 Interservice/Industry Training Simulation, and Education Conference, held 1-4 Dec 08, in Orlando FL.

14. ABSTRACT
This technical paper contains the contributions of the Air Force Research Laboratory, Human Effectiveness Directorate, Warfighter Readiness Research Division (AFRL/RHA) to the 2008 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) conference. I/ITSEC is the premiere event of its kind in the world of training, modeling, and simulation. The conference and exhibits represent the changing technologies as well as the changing training and education needs of its attendees. The 2008 conference theme was: Learn. Train. Win! The conference included multiple presentations of previously unpublished papers, as well as tutorials and special events--all selected by an extensive peer review process. This paper contains four AFRL/RHA paper presentations and the special I/ITSEC edition of the AFRL/RHA newsletter, *Fight's On*.

15. SUBJECT TERMS
Proceedings; Education and training technology; Flight simulation; Flight simulators; Instructional devices; Simulation and training devices; Simulators; System design and operation; Training; Distributed mission operations; DMO;

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON Dr Herbert H. Bell
a. REPORT UNCLASSIFIED	b. ABSTRACT UNCLASSIFIED	c. THIS PAGE UNCLASSIFIED	UNLIMITED	79	19b. TELEPHONE NUMBER *(include area code)*

This page intentionally left blank.

CONTENTS

2008 Committee Members:

Session E -5 Make it Real – Chair, Dr Tiffany Jastrzembski, AFRL/RHAT

Session H-1 Come Fly with Me – Chair: Dr Liz Gehr, The Boeing Company; Deputy Chair: Oscar Garcia, AFRL/RHAS

Session S-3 Missiles, Hardware, and Passports – Chair: Dr Glenn Gunzelmann, AFRL/RHAT

Training Interventions for Reducing Flight Mishaps

Gregg A. Montijo and David Kaiser	**V. Alan Spiker**	**Robert Nullmeyer**
Crew Training International	**Anacapa Sciences**	**Air Force Research Laboratory**
Memphis, TN	**Santa Barbara, CA**	**Mesa, AZ**
gmontijo@cti-crm.com	vaspiker@anacapasciences.com	robert.nullmeyer@af.mil
dkaiser@cti-crm.com		

ABSTRACT

Increasing numbers of preventable mishaps across all military services led Secretary Rumsfeld and all Service Chiefs to call for a reduction in such events by 75% from 2003 levels. Most were attributed to human error. The highly task-loaded training and combat missions flown by fighter pilots place particularly high demands on effective management of cockpit resources for safe and successful mission accomplishment. While every flight training program already includes some form of resource management training, there is surprisingly little evidence regarding the effectiveness of varying training approaches to reduce flight mishaps.

This paper describes a project to help the Air Force reduce preventable mishaps by determining the specific root causes of fighter and unmanned aerial system mishaps, developing behaviorally-based training objectives, identifying promising training media alternatives, and defining specific measures of effectiveness. Mishap reports revealed several repeating problems in the areas of situation awareness, task management, and decision making in all platforms studied. A Delphi Panel of fighter, attack, and Predator pilots reviewed and in some cases, amplified the specific underlying human factors that are most challenging to pilots in tactical environments. The panel also considered the feasibility and probable value of nine potential training interventions. The Predator community was chosen for implementation and assessment of four interventions – focused academic training, interactive case histories, game-based multi-task practice, and a laptop-based simulator for team training. A review of historical Predator student records revealed that many trainees have difficulty mastering attention management, task prioritization, selecting a good course of action, and crew coordination.

Spiral implementation will enable the contributions of each intervention to be assessed using a controlled experimental design at an operational training unit. Anticipated benefits include increased student situation awareness, more effective task management, and improved decision making in subsequent flights, all contributing to the ultimate goal, fewer mishaps.

ABOUT THE AUTHORS

Gregg A. Montijo is the Combat Air Forces Crew/Cockpit Resource Management Program Manager for Crew Training International, Inc. He is a retired USAF command pilot with over 4,000 hours in the A-10, F-16, A-320 aircraft and various gliders. He earned a B.S. from the U.S. Air Force Academy in 1981, specializing in Human Factors Engineering. He also earned an M.S. in Procurement and Acquisition Management from Webster University in 1995. He holds an Airline Transport Pilot's license with a type rating in the B737 and has over 10 years program management experience at the Pentagon and in the civilian sector. He has been with Crew Training International since 2001.

David Kaiser is the Director of Courseware Development for Crew Training International and is responsible for all courseware development for both Crew Resource Management and Aircrew Training. He earned his master's degree from Embry Riddle Aeronautical University with a dual specialization in Aviation Education and Aviation Safety. A Navy aviator with combat experience, he is currently active in the Navy Reserve. He has managed various experimental projects and has been personally awarded two U.S. patents.

Alan Spiker is a principal scientist at Anacapa Sciences where he has worked since 1982. He received his PhD in experimental psychology from the University of New Mexico in 1978. At Anacapa, he is responsible for managing human factors research projects in advanced technology, and has specialized in aviation training and human performance for all branches of service and the airline industry. During the past 10 years, he has conducted research

on the cognitive underpinnings of effective, safe aircraft operation, including crew resource management, multi-tasking, planning, and decision-making.

Robert Nullmeyer is a research psychologist with the Air Force Research Laboratory at Mesa, AZ, where he has been since 1981. He has conducted research on training system evaluation, simulator training effectiveness, and crew resource management training needs analysis. He earned a PhD in Experimental Psychology from the University of New Mexico in 1979.

Training Interventions for Reducing Flight Mishaps

Gregg A. Montijo and David Kaiser
Crew Training International
Memphis, TN
gmontijo@cti-crm.com
dkaiser@cti-crm.com

V. Alan Spiker
Anacapa Sciences
Santa Barbara, CA
vaspiker@anacapasciences.com

Robert Nullmeyer
Air Force Research Laboratory
Mesa, AZ
robert.nullmeyer@af.mil

INTRODUCTION

The central role of human error in flight mishaps is well documented. Helmreich and Fouchee (1993) reported that flight crew actions were <u>causal</u> in more than 70% of worldwide air carrier accidents from 1959 to 1989 involving aircraft damaged beyond repair. In commercial aviation, mishaps attributed to human error appear to be declining. Shappell, Detwiler, Holcomb, Hackworth, Boquet, and Wiegmann (2006) reported a steady decline in percentages of commercial aviation accidents in which human error was causal from 73% in the early 1990s to less than 60% in 2000-2002. Similarly, Baker, Qaing, Rebok, and Li (2008) reported a drop in air carrier mishaps involving human error from 42% in the 1980s to 25% in 1998-2002.

In contrast, mishap rates rose slightly but steadily from 1999 through 2003 in all U.S. military services following decades of improvement. In the Air Force, Luna (2001) reported that human factors were causal or major contributors in over 60% of Class A mishaps from FY1991 through FY2000. Heupel, Hughes, Musselman, and Dopslaf (2007) reported similar percentages in Air Force mishaps from FY2000-FY2006 (64%). Rising mishap rates across all military services led to directives from Secretary Rumsfeld to reduce preventable mishaps (Rumsfeld, 2003, 2006). This, in turn, generated pledges from all Service Chiefs of Staff to reduce preventable mishaps by 75% from 2003 levels. The U.S. Coast Guard (2008) compared 2007 Class A flight mishap rates across all military services. Relative to mishap rates in the preceding four years, some organizations showed more progress toward reducing mishaps than did others. The Navy and Marine Corps reduced mishap rates by about one third in 2007 relative to the previous four years. The Coast guard had no Class A mishaps in 2007. In contrast, mishap rates in 2007 increased slightly service-wide in the Air Force and Army compared to the previous four years.

Further analyses of FY2007 Air Force Class A mishaps revealed unusually high numbers of F-15 and F-16 mishaps. There were six F-15 Class A mishaps in 2007 versus 2.8 per year from FY2003-FY2006. Two were attributed to human error. Thirteen F-16 mishaps rose from 6.8 historically. Seven involved human factors. Predator mishap counts rose slightly to five in 2007 from an historic average of 4.5. Three involved

human factors. These three platforms accounted for 80% of all Air Force Class A mishaps in 2007, and half of these were attributed to human factors.

In light of enviable reductions in human factors-related commercial aviation mishaps, it may be useful to review safety training practices in that arena. Helmreich, Merritt, and Wilhelm (1999) documented a progression of crew resource management (CRM) training philosophies and goals through four distinct generations. They concluded that the original safety-related goals of CRM appeared to have become lost over time and proposed a fifth generation of CRM training explicitly focused on error management. Five data sources were recommended to sharpen that focus: (a) formal evaluations of flight crews, (b) incident reports from aviators, (c) surveys of flight crew perceptions regarding safety and human factors, (d) parameters of flight from flight data recorders, and (e) line operations safety audits (LOSA). Each illuminates a different aspect of flight operations.

Helmreich, Wilhelm, Klinect, and Merritt, (2001) studied threats to safety and the nature of errors in three airlines using LOSAs. Striking differences were observed among these air carriers regarding both threats to safety and the nature of operator errors. Based on this experience, Helmreich and his colleagues concluded that individual air carriers cannot assume their training requirements will correspond to normative data from the industry. Rather, they postulated that organizations must have current and accurate data regarding the true nature of threats and errors to shape effective training content and structure assessments of training impacts. They proposed a sixth generation of CRM training that adds the need to understand an organization's threats to safety to the previous domain of error management.

We believe that threats to safety in military operations need to be better understood and error reduction training needs to be more focused if the military is to achieve the desired reductions in preventable mishaps that have been enjoyed by their commercial counterparts. To that end, several analyses of Air Force mishap data were recently completed. Nullmeyer, Stella, Montijo and Harden (2005) analyzed attack, fighter, and tactical airlift mishaps, and Nullmeyer, Herz, Montijo and Leonik (2007) investigated Predator mishaps. Both reconnaissance

(RQ-1) and multi-mission (MQ-1) platforms were included in the Predator analyses. Three skill areas were consistently cited as factors in Air Force fighter, attack, and Predator flight mishaps: (a) situational awareness development and maintenance, (b) task management, and (c) decision making.

We recognize that mishap reports are not sufficient by themselves to structure training. Dekker (2003) described several potential problems associated with over-reliance on human error taxonomies, including risks associated with removing the context that helped produce the error. Such concerns imply that quantitative mishap human factors trends must be viewed in the context of other information to develop truly robust training interventions that are likely to impact safety and effectiveness. To that end, we augmented the safety data with expert opinion and trends in student records.

The remainder of this paper describes a project that intends to help the Air Force reduce preventable mishaps by determining the particular human factors skills that are most relevant to the fighter and unmanned aerial vehicle (UAV) communities, identifying several potential strategies for reducing subsequent operator error through training, and developing a concept of operations to test the effectiveness of the most promising training interventions that would address deficient skills.

AIR FORCE CLASS A MISHAPS

The first step in this project was to identify current human factors deficiencies in high-workload fighter and UAV tactical environments. To accomplish this, we reviewed reports of A-10, F-15C, F-15E, F-16, and RQ-1/MQ-1 Class A mishaps ($1 million damage or fatality) from FY1996 through FY2006. The Air Force Safety Center (AFSC) documents Class A mishaps in a variety of forms, and our analyses combined information from several of these sources. The first

was a detailed **Human Factors Database** populated and maintained by AFSC Life Sciences Division analysts. This database lists all factors cited regarding the roles played by operators, maintainers, and other personnel in each mishap. Our research team created aircraft-specific databases to facilitate identification of trends and idiosyncratic results. We further investigated specific causes and contributing factors using more detailed mishap source documents, primarily the **Safety Investigation Board Report** and the **Life Sciences Report**. Qualitative analyses of discussions in these reports were accomplished to gain a better understanding of the underlying behaviors that led to each element being cited.

The AFSC Human Factors Database listed all human factors cited in the Life Sciences Report section of each full mishap report and provides a contribution score: 4=causal, 3=major factor, 2=minor factor, 1=minimal factor and 0=present, but not a factor for mishaps through FY2006. From this database, we created a combined index (frequency and importance) by summing these scores across mishaps for each cited human factors element. These weighted sums were then used to rank-order the individual elements, with a separate ranking created for each weapon system. The top ten causal and major contributing factors cited in the Human Factors Database across the platforms addressed in this study are shown in Table 1. Numbers of mishaps by weapon system are listed immediately beneath each aircraft type. For example, there were 20 A-10 mishaps. The remainder of the table shows the numbers of mishaps in which a specific human factors element was cited as a causal or major factor. In the 20 A-10 mishaps, channelized attention was cited in nine mishaps, and task misprioritization was cited in seven.

Channelized attention, task misprioritization, and *selecting the wrong course of action* were cited as problems in every platform analyzed. Factors beyond these top ten were also cited, but usually in only one

Table 1. Top Ten Root Causes in Tactical Aircraft Class A Mishaps (FY1996-FY2006)

Aircraft Type (Numbers of mishaps) **Human Factors Elements:**	A-10 (20)	F-15C (14)	F-15E (9)	F-16 (86)	RQ-1/MQ-1 (30)
Channelized attention	9	8	3	25	8
Task misprioritization	7	3	2	17	4
Misperception	4	4		14	
Selecting wrong course of action	3	3	1	9	4
Wrong technique/procedure		6	1	8	4
Cognitive task oversaturation	5	3		10	
Spatial disorientation	3		2	11	
Risk assessment				11	3
Distraction/inattention			3	7	3
Inadequate in-flight analysis	7	2	1		2

or two platform types. Necessary action delayed and event proficiency were problematic in A-10 mishaps. Crew coordination, checklist error, confusion, inadequate written procedures, and interface design issues were commonly cited in Predator mishap reports. These quantitative analyses suggest that a number of threats to safety are common across fighter, attack, and reconnaissance platforms, but there are a number of platform unique issues as well, particularly for Predator operators.

CANDIDATE TRAINING INTERVENTIONS

Through reviews of Web planning-related sites, technical descriptions of interventions in the literature, and discussions with training analysts, nine promising candidate training interventions were identified that would address the skills emerging from the mishap analyses. The interventions spanned the spectrum of possible solutions from self-study and focused academics to specialized simulation and network technologies. We defined a "promising" intervention as one that has a potentially positive impact on one or more of the HF skill deficiencies identified, is logistically and technologically compatible with a mission-oriented training environment, and is feasible for implementation in this Phase II Small Business Innovative Research project, (i.e., can be implemented and evaluated within program time and budget constraints [2 years and $750,000]). The interventions were not necessarily mutually exclusive, and could, as needed, be bundled into a more comprehensive intervention "package." The nine identified candidate training interventions are shown in Table 2.

Table 2: Potential Training Interventions

Intervention	Description	Example
Self Study	Material is presented to the aircrews in text format via e-Learning to study at their own pace.	Chair Fly or Table Top a Mission - Warfighter might review choke points in a mission during pre-flight and think through courses of action that could be taken to reduce workload ahead of time.
Classroom-Style Training	Material is presented via a number of delivery styles: • Pure lecture • Guided lecture and discussions • Facilitated lecture (guided learning) • Facilitated lecture with in-class exercises • Computer-based self-study, plus facilitated advanced in-class interactive case studies/exercises	Videos could be taken of successful and unsuccessful crews performing the HF skill of interest in the mission trainer. To enhance instruction, the videos could be "scripted," using role-playing instructors, to highlight particular HF positive or negative behaviors.
Computer-Based Training	Training can be provided in specific skills, where a background scenario could be given to "draw" the warfighter into the context.	The team trainer GemaSim - Crews are given academics to understand their individual reactions to stress, how to recognize stress limits of others, and how to function effectively as a team under stress. Crews are assigned to a laptop-based network to complete a mission (space) exploration in which they compete against other teams of crewmembers. During the mission they are subjected to stress in order to experience breakdown in cognitive capabilities. Crews are observed and debriefed on their experience.
Part Task Trainer	A moderate fidelity simulator could be designed that has high fidelity for the HF skill of interest, with lower fidelity for other parts of the mission. • Specially designed equipment • Existing equipment with specific software or mission profiles	A CRM Part Task Trainer (PTT) was developed for the C-130 community that had fully functioning radios so copilots and navigators could learn to communicate during airdrops. The rest of the simulator – flight controls, visuals, multi-function display – was of lower fidelity, just enough to support the aircrew for the other parts of the mission.

<u>Gaming Solution</u>	CBT instructional material transformed into a game where points are awarded, repetitive play is encouraged, and competition is emphasized by displaying the top scores.	Game requiring players to monitor and respond to several simulations of UAV displays (e.g., heads-up display screens, chat lines, imagery, map, etc.)
<u>Full Mission Trainers</u>	Correct skill deficiencies in a Full Mission Trainer environment • Add software to existing system • Modify mission profile to train skill	Simulators can be configured that have fairly high fidelity to support multi-crew teamwork training in customized scenarios. Problem HF skills can be addressed through repetitive practice, feedback, and debriefs.
<u>Dedicated Mission Trainers</u>	Simulator training specifically dedicated to specific skills tied to safety of flight • Use existing simulators and modify software to train specific CRM skills • Relies heavily on debriefing	Simulators that emphasize particular missions can be used where the targeted HF skills are a major player for that mission. (e.g., channelized Attention could be selected for highlighting training in the context of air/ground missions with visually complex enemy laydowns).
<u>Modify Existing Simulator Profiles</u>	Use existing training capabilities, insert specific training events that would stress and target particular HF skills. • Requires in-depth analysis of existing profiles • Specific mission events are needed to have desired behavioral outcomes • "The Gouge" can quickly develop among flight crew - negates training • Easiest in terms of schedule, cost	A particular training profile could be modified by inserting additional task stressors, (e.g., threat pop-ups, reduced visibility, caution lights, etc.), to provide training in task prioritization. Embedded performance standards would be included in the events, as well as feedback provided in the debrief.
<u>Networked Solutions</u>	Full spectrum missions flown in simulators linked with other participants • May be stand-alone in nature or part of a Joint exercise. • May blend real world and synthetic environments. • The ability to capture individual behavior in a dynamic computer environment with a wide-range of possible outcomes is a potential challenge	Distributed Mission Training (DMT)/ Distributed Mission Operations (DMO)

THE DELPHI PANEL

A Delphi Panel of F-15, F-16, A-10, and RQ-1/MQ-1 warfighter experts was convened to solicit their opinions on skill deficiencies and potential training interventions. To accomplish these goals, we constructed a multi-faceted instrument designed to collect both quantitative data regarding problem frequency and difficulty, and qualitative data reflecting the panel's comments regarding key problems, issues, and explanations. As such, the instrument was consistent with the project's multi-method, multi-measure approach to identifying, defining, measuring, and evaluating high-payoff CRM skills. Because of high Operations Tempo (OPTEMPO) and scheduling issues, we restricted our

panel to a half-day at the U.S. Air Force Weapons School, Nellis AFB, NV. This location permitted at least one representative from each of the aforementioned weapon systems to attend, with the Predator community supplying three people. Thus, a total of six experts attended the three-hour session. Despite the logistical problems in convening the panel, the qualifications and experience levels of the participants were impressive. All were officers, O-4 and above, with most having hundreds or thousands of hours operational training and combat experience with their particular weapon system. All participants were highly-motivated to support the present project, and each appeared to be genuinely interested in improving CRM skills for their weapon system. In short, the

panel composition and tone was ideal for our purposes.

Identifying Human Factors Skills

Panel members were given a list of skills that had been derived from the Class A mishap reports. The list included 19 skills – the ten factors listed in Table 1 plus nine others. In each case, panel members were asked to rate each skill using the following five-point scale:

1. No problems in training/operational missions
2. Minor problems in training/operational missions
3. Some problems in training/operational missions
4. Major problems in training/operational missions
5. Severe problems in training/operational missions

Panel participants reviewed each skill in turn, providing a rating and, in some cases, offering written comments explaining the basis for their ratings. A moderated discussion concerning issues and problems regarding these skills followed.

The initial series of analyses was performed on the data from the six panelists who represented all four tactical weapons systems (three of the six panelists were Predator operators). Table 3 summarizes the mean importance/problem ratings for the skills that were identified in the mishap report analyses. They are presented in descending order of mean rating, where the scale can range from 5 (severe problem) to 1 (no problem). The top four skills based on mishap reports are indicated in red italics. To provide a metric for

making comparisons, we computed the variance of ratings within each skill, took the average, and then computed the average standard error about the mean. Doubling that number provides a good estimate of the typical rating difference that would be considered statistically significant if inferential tests were conducted (Hays, 1973). Our analysis showed this value to be about .75. For example, on the basis of this metric, we could conclude that the average rating for Cognitive Task Oversaturation (3.7) is statistically higher than Task Misprioritization (2.9). While not used to completely guide our analyses or interpretations, such an index should be kept in mind when attempting to draw firm conclusions from an admittedly small sample size.

The quality of the information provided, given the high experience levels of the panelists, more than compensates for the lack of statistical power in any test that one would conduct. It is evident from the table that although the top four human factors topics from mishap trends are, by and large, among the higher-rated problems, there are others that the experts elevated in terms of relative importance. In particular, Cognitive Task Oversaturation was the factor that was rated as being most problematic by the Delphi Panel, even though it did not occupy that spot in any platform based on mishap report analyses, and was not cited at all in Predator mishap reports. This element refers to the magnitude or variety of inputs exceeding operator limitations to process information.

Table 3. Mean Rating of Importance/Problem for 19 Human Factors

Human Factor	**Mean Rating** (5=max, 1=min)
Cognitive Task Oversaturation	3.7
Channelized Attention	*3.4*
Inadvertent Operation	3.3
Inadequate In-flight Analysis	3.0
Confusion	3.0
Wrong Course of Action Selected	*3.0*
Task Misprioritization	*2.9*
Crew Coordination Breakdown	2.9
Misperception of Speed, Distance, Altitude	*2.8*
Wrong Technique	2.6
Distraction	2.5
Limited Systems Knowledge	2.4
Poor Intracockpit Communication	2.4
Checklist Error	2.3
Inattention	2.2
Complacency	2.2
Subordinate Style	2.0
Overcommitment	2.0
Poor Risk Assessment	1.8

Note: All four tactical weapon systems are included.

Inadvertent Operation reflects a poor choice of switch or function operation, which is especially problematic with the software intensive Predator operator console. Inadequate Inflight Analysis and Confusion are problem areas that appear as factors in multiple systems.

Selecting Training Interventions

The Delphi session then turned to candidate training interventions. The research team explained the nine different training interventions the panel would be asked to consider, corresponding to the ones listed in Table 2. The interventions were presented in reverse order of fidelity, beginning with self-study, followed by classroom-style training, computer-based solutions, full mission trainers, dedicated mission trainers, modification of existing simulator profiles, and networked solutions. Note that these interventions are actually *categories* of technologies that span a spectrum of possible solutions to the HF skills problems provided in the first part of the Delphi session. The presentation was interactive, with panel members asking questions and offering suggestions. Two ratings were asked of each of the nine interventions. The first was a five-point, behaviorally-anchored scale that had participants rate the intervention's estimated degree of impact on the targeted human factors skills. A second five-point scale called for rating the feasibility of implementing the intervention in an operational training squadron. Besides the rating, the instruments contained space for panel members to make amplifying comments; free-flowing discussions followed the rating process.

During the Delphi Panel session, one of the panel members, the commander of the 11th Reconnaissance Squadron (RS), indicated his desire to have other members of his squadron review the instrument and provide their assessment. The commander's endorsement of the project, and his willingness to have the MQ-1 Predator community serve as claimants, was unquestionably a turning point in the project. Per the commander's suggestions, we supplied the squadron with additional copies of the instruments. Several weeks after the workshop, three additional completed instruments were provided to the project team. It was at this point that we decided to perform two analyses. The first was on data from the six original Delphi Panel members. The second was on the six MQ-1 operators, three from the Delphi session and three survey respondents from the 11th RS, who comprised our sample. The SMEs from the other platforms provided highly similar ratings, so only the ratings from the six MQ-1 operators are shown in Table 4. The left part of the table summarizes the mean ratings of expected impact in descending order; the right portion provides the average ratings for intervention feasibility.

As can be seen, there is a marked divergence between the two sets of ratings. The interventions that panel participants rated as having the highest impact were mostly associated with being the least feasible to implement, and vice versa. Analysis of the comment data provides some ready explanations for these results. In this regard, full mission trainers were clearly seen as an effective way to train many human factors skills. Unfortunately, their feasibility for implementation within the time and resource constraints of this project is limited. Conversely, computer-based training, which was summarily dismissed by attendees based on recent negative experience, was rated poorest on impact yet was recognized for being quite feasible. It should be noted that classroom training, the clear favorite for feasibility, also received respectable marks for potential impact. This bodes well for attempts to improve error reduction via classroom training by targeting specific human factors skills with new case examples and highly focused spin-up training. This issue is taken up later in the paper when we discuss the interventions chosen for implementation.

Table 4. Mean Ratings of Intervention Impact and Feasibility (RQ-1/MQ-1 only)

Intervention Impact		Intervention Feasibility	
Intervention	Mean Rating	Intervention	Mean Rating
Full Mission Trainer	4.3	Classroom Training	3.8
Classroom	4.2	Computer Based Training	3.3
Dedicated Mission Trainer	4.1	Handheld Game	3.3
Modify Existing Simulator	3.8	Self Study	3.2
Self Study	3.6	Network Solutions	3.0
Part Task Trainer	3.5	Part Task Trainer	2.5
Network Solutions	3.2	Full Mission Trainer	2.5
Handheld Game	3.0	Dedicated Mission Trainer	2.4
Computer-Based Training	2.7	Modify Existing Simulator	2.2

Finally, we received the endorsement of the 11th RS Commander to host field studies of resulting training interventions. Having an operational claimant who eagerly awaits our interventions ("I would like them today!") is a reaction that is all-too-rare in the research and development community. As we describe below, we plan to work extremely closely with the 11th RS Commander and his organization to ensure that the training interventions we specify, prototype, develop, and implement meet the squadron's current and projected training requirements.

TRAINING RECORDS ANALYSIS

With the selection of the Predator training program as the environment in which interventions would be implemented and evaluated, training records in this community were analyzed to identify tasks that are particularly difficult or challenging for students, conducting both quantitative analyses on grades and content analyses on instructor comments.

Records from 70 student pilots and 75 sensor operators were reviewed from the Predator Operator Basic and Requalification course, focusing on student performance in the final 2 flying training sessions preceding the checkride. Instructors used a 5-point grading scale from 0 to 4, with a "2" or higher representing a passing level of performance. No "0" scores were observed, but 101 "1s" were recorded for pilots and 62 "1s" for sensor operators. These less-than-passing grades at the end of training were concentrated in 7 of the 45 graded pilot task elements and 4 of the 50 sensor operator task elements.

For pilots, the task elements were:

- Buddy laze procedures
- Launch
- Target acquisition, aircraft position
- Operational mission procedures
- Deconfliction plan/execution
- AGM-114 employment
- Airmanship/aircraft control

For Sensor operators, the task elements were:

- Launch
- Mission CRM/crew coordination
- Mission planning/preparation
- AGM-114 employment

These problematic task elements were further analyzed with the aid of instructors to identify common underlying skill areas. Four skill areas emerged: avoiding channelized attention, Prioritizing tasks, selecting an appropriate course of action, and crew coordination. Two particularly challenging syllabus events were also identified that require students to apply these skills: a simulator-based emergency procedures scenario, and a flightline tactical mission that occurs shortly before the final checkride.

TRAINING INTERVENTIONS SELECTED

To accelerate skill development in the problem areas that emerged from the preceding activities, four training interventions were selected for further development and evaluation: enhanced focus academic training; interactive, web-based or desktop case histories; gaming computer-based training to develop individual task monitoring and task management skills; and a computer-based team training environment. Each is further described below.

Enhanced focus academic training is based on the foundations of adult learning principles. These principles are presented in a facilitation style, in contrast to lecture style, in order to actively engage the following androgological principles (Knowles 1980; Knowles, Holton & Swanson 1998): (a) fulfilling the learner's need to know (helping students see the value of training and how it applies to them in their job); (b) allowing students to be more self-directed; (c) leveraging a variety of experiences to build on some learners' already-acquired experiences, transferring that knowledge base to those who have less experience; and (d) specifically designing the learners' experience to increase their readiness, orientation, and motivation to learn.

Interactive, web-based or desktop case history is based on a computer-based training system developed for the Navy that took articles from the Navy's *Approach* magazine, added supplemental information to reinforce core concepts in human performance disciplines, and presented this information in electronic form (Spiker, Hunt, and Walls, 2005). It was intended for use as an adjunct to classroom instruction. The summaries are written in a readable style designed to both entertain and educate. The case study is followed by a short set of fairly difficult questions that are written to require the student to read the case study and understand the main points. It was clear from the Delphi Panel that our experts all had less-than-stellar experiences with CBT in the past. The prevailing view was that much of what they had experienced was merely "electronic page turning," and not particularly engaging. In recognition of this, the intent with this medium is to develop compelling, interesting, informative, and memorable instruction by design.

Computer-based gaming of individual skills as an intervention is loosely adapted from a test of multi-tasking ability called *SYNWIN* (Elsmore, 1994). While SYNWIN's prior use has been as a selection test, our

plan calls for casting the concept in a game format that can be played by trainees while they are receiving their initial CRM training. Our belief is that promoting the instructional material in the form of game, where scores can be competitively acquired and even posted, will overcome some of the negative reaction to CBT that was discussed in the previous task. The test requires users to simultaneously monitor four quadrants of the primary display screen. The upper left quadrant of the screen displays a letter recall task in which participants click a button to indicate whether a probe letter was a member of a previously displayed set of letters (the subject must remember that set of letters). The upper right quadrant presents an arithmetic task, where participants solve simple, randomly-generated three-digit addition problems. A visual monitoring task is in the lower left, where participants click on a gauge to reset a slowly moving pointer before it reaches the zero mark. The lower right quadrant has an auditory monitoring task where participants listen to a series of high and low frequency tones, and click a button when they hear a high frequency tone.

From an instructional perspective, one of the strongest features of games is that they offer ample opportunity for practice and repetition. As well, games usually provide immediate, clear feedback and require criterion skill mastery to move to the next level. But the most-cited advantage of using game elements in instruction is the motivational factor – people usually want to play games and will voluntarily devote a great deal of time to mastering the skills and rules of the game. This may be particularly relevant with many of today's students and trainees who, as digital natives, have been raised in a technology-dominated environment, with hours of video and computer game playing.

Besides transforming the SYNWIN test concept into a game, we will also explore altering each of the four

tasks so they have more in common with tasks that UAV operators presently perform. For example, the memory recall task, which in SYNWIN consists of random letter/number strings, can be converted into a more meaningful task where the aviator is to recall sequences of letters and numbers that might correspond to airfield designations, waypoints, landmarks, navigation aids, etc. While the cognitive task – holding information in memory for an extended time – is the same, the actual task will more resemble what is actually required of Predator pilots and sensor operators. Similarly, the addition task could be expanded to include other mental operations that UAV operators must perform, such as doing basic geometry to compute descent angles, calculating distance between waypoints, or extrapolating airspeeds and leg times, to name a few. Similarly, the visual monitoring task does not have to be restricted to a fuel gauge. It too can be altered to more closely mimic UAV operations. For example, we could use an embedded video (say, from a sensor) and ask the subject to monitor it for some dynamic characteristic (e.g., a target).

Computer-based team training is designed to exercise team functions and behavior in a stressful environment. The GemaSim team trainer (Figure 1) allows for the experience, observation, analysis, modification and consolidation of authentic behavioral patterns that emerge under stressful conditions. Once under stress, humans may switch from established norms, industry practice, etc. and apply a different set of dominant logic pathways, resulting in abnormal behaviors. This effect has been observed in such high-risk/high-pressure industries as aviation, rail, medicine and executive management. The intent of this device is analogous to the high altitude chamber training where pilots, although taught the effects of hypoxia, all experience different symptoms. Similarly GemaSim provides an enjoyable, but serious and relevant simulation activity that allows for one's own

Figure 1: **Students under stress during GemaSim team training**

behavioral patterns to be experienced, together with those of a specific team under situations of increased pressure. Through an understanding of the causal factors of human behavior, and by analysis of one's own behavioral patterns, these can be modified, re-exercised and consolidated.

IMPACT ASSESSMENT

Our plans call for conducting an 18-month assessment of the four training interventions at Creech AFB. We plan to follow Kirkpatrick's (1996) four-level evaluation approach in which data are collected to assess: (a) the reaction of trainees to the **usability and usefulness** of the training intervention (Level I); (b) the **amount of learning or skill acquisition that occurs** from the training (Level II); (c) **if the skills that are trained transfer to the job** (flight) environment (Level III); and (d) the **benefits that accrue to the organization** as a result of the training (Level IV).

As Salas and his colleagues have noted (Salas, Fowlkes, Stout, Milanovich, & Prince, 1999), few studies of the overall effectiveness of CRM training (Level III) have been conducted, and even fewer assessed all four levels in the same context. We plan to fill this empirical data gap by implementing a series of measures at various points in the training curriculum, including a baseline period before the four interventions are introduced. A new class of pilot and sensor operator training is offered roughly every 3 weeks at the squadron, with some 20 students attending per class. Importantly, we will be performing a fairly controlled evaluation as only half the classes will receive the training interventions, with the other half serving as a control (receiving only traditional CRM). The large sample size should give us sufficient statistical power to perform multivariate analysis of variance and follow-up test procedures.

Our training interventions will be incorporated into the current curriculum as a series of four "spirals" in order to restrict our footprint on on-going operations and to help manage the complexities of parallel development. The first spiral will consist of only the first intervention (focused academics). The second spiral will entail implementing focused academics and interactive case histories. Spiral 3 will consist of the first two interventions plus the game-based training. The final spiral will comprise all four interventions. Each spiral will be implemented in two classes (about 40 students per condition), where another two classes will serve as a control. This design will let us gauge both the training effectiveness of the overall intervention package (relative to current CRM training), as well as the contributions of the individual interventions to effectiveness.

To measure intervention impact, we will employ a cadre of specialized instruments and review the squadron's regular training records. First, we will insert questions into the end-of-course critique to assess student reaction to the training in the four HF skills of interest (Level I assessment). Second, we will conclude each intervention with a comprehension assessment to ensure that learning of the HF skills has occurred (Level II).

Instructors and observers will use a specialized gradesheet to measure proficiency in the simulator training sessions following the interventions. These sessions will give us the much-needed Level III data to gauge whether the skills we believe students have learned in our training interventions actually manifest themselves in realistic flight conditions. This gradesheet will consist of some half-dozen key behaviors associated with each HF skill. For example, the HF skill "avoids channelized attention" would be decomposed into such key behaviors as: effective cross-check includes all relevant displays; cross-check does not stagnate; switches attention as the situation priority changes; etc. Importantly, key behaviors will be defined to support reliable observation by instructors and raters.

CONCLUSION

Our main purpose in this project is to help reduce preventable flight mishaps, so our assessment of benefits to the organization needs to address the impact of these interventions on safety of flight. A direct assessment of that effect will require longitudinal tracking of Predator crews beyond the time frame of this project. This project will, however, determine the ability of our interventions to accelerate the development of skills that were lacking in previous Class A mishaps.

Much of what we learned to date is encouraging. The vast majority of Air Force Class A mishaps (78%) in 2007 involved F-15, F-16, and Predator operations, and the root causes of mishaps in these three platforms have much in common—mishap reports from all three communities frequently cite channelized attention, task misprioritization, and course of action selected. Our panel of experts from each of these systems added cognitive task oversaturation as a fourth problem area. As a result, it appears that a finite set of factors is driving Air Force preventable Class A mishaps.

Our approach assumes that these problem areas reflect trainable skills. Given the support that we enjoy with the Predator community, this project represents an excellent opportunity to move from problem statements to validated solutions. Interventions that positively impact on subsequent attention and task

management or improved decision making for Predator crews should be directly applicable to the fighter and attack communities.

REFERENCES

Baker, S.P., Qiang, Y., Rebok, G.W., and Li, G. (2008). Pilot error in air carrier mishaps: longitudinal trends among 558 reports, 1982-2002. *Aviation, Space, and Environmental Medicine*, 79(1), 2-6.

Dekker, S.W.A. (2003). Illusions of explanation: A critical essay on error classification. *The International Journal of Aviation Psychology, 13*(2), 95-106.

Elsmore, T. (1994). SYNWORK1: A PC-based tool for assessment of performance in a simulated work environment. *Behavior Research Methods, Instruments & Computers, 26,* 421-426.

Hays, W.L. (1973). *Statistics for the social sciences* (2nd ed.). New York: Holt, Rinehart, & Winston.

Helmreich, R.L. & Foushee, H.C. (1993). Why crew resource management? Empirical and theoretical bases of human factors training in aviation. In E.L. Wiener, B.G. Kanki, and R.L. Helmreich (Eds.), *Cockpit resource management*, San Diego, CA: Academic Press.

Helmreich, R.L., Merritt, AC., & Wilhelm, J.A. (1999). The evolution of crew resource management training. *The International Journal of Aviation Psychology, 9*(1), 19-32.

Helmreich, R.L., Wilhelm, J.A., Klinect, J.R., & Merritt, A.C. (2001). Culture, error, and crew resource management. In E. Salas, C.A. Bowers, & E.Edens (Eds.), *Applying resource management in organizations: A guide for professionals. Hillsdale,* NJ: Erlbaum.

Heupel, K. A., Hughes, T.G., Musselman, B.T., & Dopslaf, E.R. (2007). *USAF aviation safety: FY 2006 in review*. http://safety.kirtland.af.mil

Kirkpatrick, D.L. (1996, January). Techniques for evaluating training programs. *Training & Development.*

Knowles, M.S. (1980). *The Modern Practice of Adult Education: From Andragogy to Pedagogy.* Englewood Cliffs, NJ: Cambridge Adult Education.

Knowles, M.S., Holton, E.F., III, &. Swanson, R.A. (1998). *The Adult Learner.* Houston: Gulf Publishing.

Luna, T.D. (2001). *USAF aviation safety: FY 2000 in review*. http://safety.kirtland.af.mil

Nullmeyer, R.T., Stella, D., Montijo, G.A., & Harden, S.W. (2005). Human factors in Air Force flight mishaps: Implications for change. *Proceedings of the 26th Interservice/Industry Training Systems and Education Conference.* Orlando, FL.

Nullmeyer, R.T., Herz, R.A., Montijo, G.A., & Leonik, R. (2007) Birds of Prey: Training solutions to human factors problems. *Proceedings of the 28th Interservice/Industry Training Systems and Education Conference.* Orlando, FL.

Rumsfeld, D. (2003). *Reducing Preventable Accidents.* Memorandum for Secretaries of the Military Departments. The Secretary of Defense, Washington DC.

Rumsfeld, D. (2006). *Reducing Preventable Accidents.* Memorandum for Secretaries of the Military Departments. The Secretary of Defense, Washington DC.

Salas, E., Fowlkes, J.E., Stout, R.J., Milanovich, D.M., & Prince, C. (1999). Does CRM training improve teamwork skills in the cockpit? Two evaluation studies. *Human Factors*, *41*(2), 326-343.

Shappell, S., Detwiler, C., Holcomb, K., Hackworth, C., Boquet, A., & Wiegmann, D. (2006). Human error and commercial aviation accidents: A comprehensive, fine-grained analysis using HFACS (DOT/FAA/AM-06/18). Washington, D.C.

Spiker, V.A., Hunt, S.K., & Walls, W.F. (2005). *User reaction to annotated approach* (CNAP Technical Memorandum). San Diego, CA: Commander Naval Air Force, US Pacific Fleet.

U.S. Coast Guard (2008). *FY07aviation safety report,* available at http://www.uscg.mil/safety.

Computer Generated Forces for Joint Close Air Support and Live Virtual Constructive Training

Mr. Craig Eidman
Air Force Research Laboratory, Warfighter
Readiness Research Division, Mesa, AZ
craig.eidman@mesa.afmc.af.mil

1Lt Clinton Kam
Air Force Research Laboratory, Warfighter
Readiness Research Division, Mesa, AZ
clinton.kam@mesa.afmc.af.mil

ABSTRACT

Conducting robust, reoccurring Joint CAS training for Terminal Attack Controllers (JTACs) on live ranges is problematic. While stationary observation points and targets are useful for initial and basic call for fire training, live bombing ranges do not provide mobile, realistic targets for training in troops in contact, joint/coalition training, and operations in urban terrain. Distributed simulation and Live-Virtual-Constructive networks can provide JTACS with training to enhance their team, inter-team, and joint skills with greater frequency, at lower cost, and potentially more combat realism than live-range training exercises. One of the key advantages of distributed simulation training for JTACs working with attack aircraft, is that the activities can be focused on specific skills such preparing and communicating 9-line coordination briefings, procedurally "talking aircraft on to" targets, and coordinating for directives, priorities and deconfliction of fires. Fidelity requirements for computer generated forces (CGFs) have typically revolved around air-to-air fighter training or large scale wargaming. In 2004, the Air Force Research Laboratory initiated a Joint Terminal Attack Control Training and Rehearsal System research and development project. The goal of this effort was enhancing JTAC readiness by designing, developing and evaluating an immersive, DMO compatible training system using fully integrated JTAC equipment. After initial system evaluations by JTAC subject matter experts, it was apparent that the CGF scripting, intelligent behavior, systems models, and weapons would need major modifications to support effective JCAS training. To overcome these difficulties researchers developed a rapidly customizable CGF environment and instructor operator station. This paper discusses some of the unique modifications made to CGFs to support JTAC training and overall lessons learned from modeling and simulation of the JTAC environment to include behavior scripting, artillery models, realistic air-to-ground weapons delivery simulation, modeling the air-to-ground C2 environment, instructor tools, and scenario management.

ABOUT THE AUTHORS

Mr. Craig Eidman is an Electrical Engineer with the Air Force Research Laboratory, Warfighter Readiness Research Division, Mesa Research Site. He is the Lead Engineer for all laboratory synthetic environments, Electronic Warfare Training, and advanced targeting pod S&T efforts. Mr. Eidman attended the United States Air Force Academy where he earned a Bachelor of Science in Electrical Engineering in 1983. He served in the Air Force as a Command Pilot flying F-16, A-10, and OA-37 aircraft in multiple theaters and as an Evaluator Forward Air Controller. Additionally, he has a Masters of Aeronautical Science, Embry-Riddle Aeronautical University, With Distinction and is an Outstanding Graduate of the Air War College.

1Lt Clinton Kam is currently assigned to the Warfighter Readiness Research Division in Mesa, Arizona working as the Threat Systems Engineer for the Immersive Environments Branch. Lt Kam holds a B.S. in Aerospace Engineering from the University of Texas. He has worked on several projects spanning the modeling and simulation and training spectrum including simulation performance testing and optimization, HLA networking, and high fidelity piloted and unpiloted threat models.

Computer Generated Forces for Joint Close Air Support and Live Virtual Constructive Training

Mr. Craig Eidman
Air Force Research Laboratory, Warfighter
Readiness Research Division, Mesa, AZ
craig.eidman@meas.afmc.af.mil

1Lt Clinton Kam
Air Force Research Laboratory, Warfighter
Readiness Research Division, Mesa, AZ
clinton.kam@meas.afmc.af.mil

JCAS TRAINING REQUIREMENTS

Conducting robust, reoccurring Joint Close Air Support (JCAS) training for Joint Terminal Attack Controllers (JTACs) on live ranges is challenging. While fixed observation points and stationary targets are useful for initial and basic call for fire training, live bombing ranges do not provide mobile, realistic targets for training in troops in contact, joint/coalition integration, airspace deconfliction, operations in urban terrain and advanced tactics development.

JTAC Live Range Training Shortfalls

For a JTAC, the live fire training range environment is often a limited representation of actual combat operations. A typical airstrike control training event on a live range may have a small JTAC team operating independently at a pre-surveyed observation position, coordinating with a single 2-ship of attack aircraft engaging various mock-up targets with either training munitions (if allowed) or more likely "dry passes" where weapons deliveries are notional. Range target arrays are typically maximized for aircrew training and not JCAS training (often airfield complexes). If live ordnance is used, it is only on specific targets, often miles away from the JTAC location. Any realistic coordination with ground forces, artillery fires, and moving targets does not occur. Troops in contact can only be done in a "notional" sense – real ordnance or even training ordnance cannot be expended in the vicinity of the ground parties for safety reasons.

Compare this with a JTAC in a fully joint exercise or actual combat. Enemy targets are mobile, hidden, and exposed for only a limited amount of time. The JTAC is coordinating through three to four different radio networks simultaneously to control fighters, manage airspace, coordinate with ground units and deconflict fires. The observation point for an airstrike may not be optimal, in fact the JTAC may not even have "eyes on target". Intentional and unintentional obscurants or weather may hamper vision. In a worst case scenario, troops will be engaged in actual fire fights at close distances.

Scheduling and range availability are also limiting factors. In the majority of cases, JTACS are assigned with US Army units and may not be close to impact areas or ranges used by live aircraft. On many of these Army ranges the target arrays are designed for ground operations and not air operations. JTACS must travel to Air Force ranges requiring coordinated scheduling and the transport of tactical equipment to practice live call for fire training. Operational pace for both the JTAC units and the supporting attack aircraft units make this coordination challenging.

The costs in fuel, travel and equipment wear and tear are a burden to many operational units. Quite often live fire range training entails only the use of portable battery powered radios due to the limited availability and cost of vehicle mounted radio pallets. Other critical systems necessary in combat may also be unavailable. For example, JTACS in Opertion Iraqi Freedom and Operation Enduring Freedom regularly employ systems like the Remote-Operations Video-Enhanced Receiver (ROVER) to conduct airstrikes. This system receives streaming data from airborne sensor platforms like Unmanned Aerial vehicles (UAVs) or fighter and bomber aircraft targeting pods. (Erwin, 2008) The supporting sensor platforms are often unavailable for training activities. (USAF, 2007)

Finally, the Air Force centric range is often a poor representation of the joint or coalition combat environment. In a true joint environment a JTAC is managing airspace, deconflicting indirect fires, managing joint suppression of enemy air defenses, coordinating with the ground forces chain of command and fire centers and coordinating with the air support operations center (ASOC), all while controlling the actual airstrike. None of these complex tasks are available on most Air Force bombing ranges unless other Tactical Air Control Party (TACP) members role play these agencies.

These training shortfalls are well understood by senior policy officials. According to a 2002 United States General Accounting Office report on issues relating to training and equipment issues hampering air support to ground units:

"We found that adequate realistic training is often not available because of (1) Ground and air forces have limited opportunities to train together in a joint environment. When such joint training does occur, according to DOD reports and unit officials, it is often ineffective. (2) Similarly, the training that troops receive at their home stations is usually unrealistic because of range restrictions; moreover, it lacks variety—for example, pilots often receive rote, repetitive training because of limited air space and other restrictions." (United States Government Accounting Office, 2003)

Simulation for JTAC Training

Distributed simulation and Live-Virtual-Constructive networks can provide JTACS with training to enhance their team, inter-team and joint skills with greater frequency, at lower cost and potentially more combat realism than live-range training exercises. One of the key advantages of distributed simulation training for JTACs working with attack aircraft is that the activities can be focused on specific skills such as preparing and communicating 9-line coordination briefings, procedurally "talking aircraft on to" targets, coordinating for directives, priorities and deconfliction of fires. The 2007 Joint Close Air Support Action Plan recognizes that simulation now offers realistic and affordable training options to compensate for these gaps:

"Although simulation will never replace all live JCAS training, current technology allows credible substitution for specific events in initial, continuation and collective training for air and ground personnel and units. Stand-alone virtual simulators may enhance training opportunities and potentially mitigate the shortfall in selected JTAC training events for initial qualification and continuation training. Current Service, USJFCOM and USSOCOM efforts already contain many foundation elements for virtual collective training. Constructive simulations that network staff and liaison elements to practice battle management and fire support integration are also feasible." (JCAS Action Plan, 2007)

Simulation also enables advanced training and tactics development and validation. The success of Distributed Mission Operations for air-to-air training is an example of this success. During current ground conflicts, new systems, missions and weapons platforms have been integrated into the JCAS environment utilizing un-practiced employment tactics. For example, in the past JCAS was limited to a subset of fighter and special operations aircraft. Today, bomber aircrews and UAVs regularly conduct precision airstrikes against targets in support of ground forces. Often the JTAC is coordinating these airstrikes from locations where he cannot observe the actual targets, yet the targets are close to friendly ground troops. Simulation allows a safe, effective methodology to develop procedures for complex tactics and troops in contact scenarios.

JCAS TRAINING RESEARCH PROGRAM

In 2004, the Air Force Research Laboratory initiated a Joint Terminal Attack Control Training and Rehearsal System (JTAC TRS) research and development project. The goal of this effort was enhancing JTAC readiness by designing, developing and evaluating an immersive, DMO compatible training system using fully integrated JTAC sensor, target designation and communications equipment operating in real time.

Part-Task JCAS Training Solutions

Acting upon an initial request from JTAC training units, AFRL worked with industry to develop a demonstration JCAS training system using a generic pilot station integrated with a single screen visualization capability for target viewing. The resulting system, the Indirect Fire-Forward-Air Control Trainer (I-FACT) was deployed at the Air Ground Operation School (AGOS) at Nellis AFB for evaluation. This successful training system has since been deployed at a variety of JTAC and Special Operations units. (Kauchak, 2008) It has proven extremely useful in basic training of JTACS to prepare and present 9-line briefings for pilots and conduct basic airstrike control interactions.

AFRL found that while these part task training solutions provide valuable training, this training was limited in scope due the fidelity of supporting models and interfaces. I-FACT was a training solution focused solely on the JTAC and his control of CAS and artillery assets and gave operators the capability of being on a simulated battlefield with appropriate ground threats and air assets. AFRL's initial system had no scripting capability for robust Computer Generated Forces (CGFs). Aircraft on CAS attacks could be created and fly only after a mission was executed. They had no orbit or ingress points, only a final attack heading for the target. The student would call in an attack heading and look for the aircraft to "Clear Hot" but at the end of the mission the aircraft would fly out to the horizon then disappear from the simulation. Similarly, artillery

models did not use physics based calculations to determine max altitudes and time of flights of their rounds. The instructor selected the location of the detonation and immediately upon execution the rounds impacted. The man in the loop flight simulation station, which played a single aircraft, did not represent the complexities of controlling multiple fighters in a single flight and multiple flights of aircraft simultaneously. The navigation and target acquisition problems faced by a real pilot in the JCAS environment were not replicated and consequently the methods and "target talk on" a JTAC would use with real aircraft were not realistic. The system operated only at an unclassified level making integration with high fidelity classified flight simulators difficult.

Fully Immersive JTAC Training Systems

To study the benefits of a more immersive training environment for JTACS, Air Force Research Laboratory (AFRL) developed a science and technology proof-of-concept Training and Rehearsal System (TRS) to provide high-fidelity, fully immersive, realistic training with real-time sensor, simulator and database correlation along with a robust instructor operator station (IOS) and scenario generation capability. This system was designed to support performance assessment of JTAC personnel as well as study technology requirements for future immersive JCAS training systems. The design would allow stand alone training driven by the IOS aided by constructive simulations as well as distributed training with other high fidelity simulators using established Distributed Interactive Simulation (DIS) protocols.

Figure 1. Fixed 360x180 FOV Dome

A visualization of an immersive ground combat environment has significantly different requirements than that of a typical flight simulator. AFRL constructed a fixed 360x180 field of view visual dome at its facility in Mesa, Arizona to initiate research

studies into immersive JTAC training. The system was developed using state-of-the-art image generators (IGs), high resolution color photo-specific databases (some sampled at as low as 40 cm) and proven system hardware. The IGs and network interfaces were identical to fielded A-10 simulators allowing shared correlated databases, 3-dimensional models, special effects and Instructor Operator control. This allowed near perfect interaction and correlation with operational A-10 units, a natural networked training audience for training research activities.

The dome's visual system was accompanied by a set of sensor devices and emulators to further immerse the student into the scenario. These devices include a simulated M-22 Binoculars, GLID II Laser Target Designator and Mk VII Laser Range Finder. In addition to the simulated devices, software was developed to give students the ability to use their actual AN/PSN -11 or 13 GPS receivers and AN/PRC-117 or PRC-148 radios.

Figure 2. Sensor Devices

The first unit deployed JTAC TRS dome was installed at the Air to Ground Operations School (AGOS) at Nellis AFB in January, 2008. Substantial feedback has been received from the schoolhouse since that time and AFRL continues its work on the JTAC program to improve the training capabilities for the students.

Computer Generated Forces

To manage the training scenarios and provide constructive models and computer generated forces, AFRL turned to their in-house CGF development platform, XCITE, to fill the role. XCITE is AFRL's prototype CGF software based on the Next Generation

Threat System (NGTS). XCITE's government owned source code can be rapidly modified to meet the requirements of various research projects. After initial system evaluations by JTAC subject matter experts, it was apparent that the CGF scripting, intelligent behavior, systems models and weapons would need major modifications to support effective JCAS training. To overcome these difficulties researchers developed a rapidly customizable CGF environment and instructor operator station.
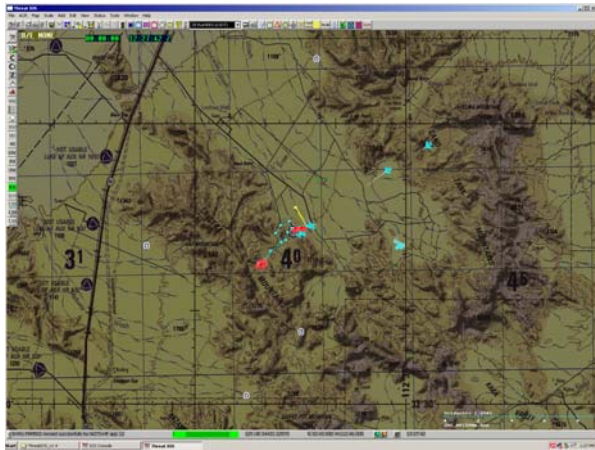

Figure 3. XCITE Instructor Operator Station

CGF SHORTFALLS AND IMPROVEMENTS

Fidelity requirements for CGFs have typically revolved around air-to-air fighter training or large scale wargaming. Initial NGTS research and design revolved around methods to conduct high fidelity, physics-based electronic warfare and air-to-air training in fighter simulators. To support this research, NGTS was designed to utilize physics-based maneuvering and aero models and high fidelity threat avionics models running at real time. Although an excellent air-to-air trainer for pilots, it did not have the capabilities for a "ground perspective" for scenario management and control. Few ground entities were modeled – mostly Surface-to-Air (SAM) sites and their associated radars. Also, the autonomous air assets had no close air support relevant tactics. New JCAS specific aircraft maneuvers, ground entities and artillery control would need to be added.

Weapons, Aircraft, and Ground Forces Models

While many aircraft air-to-ground weapons models were available in XCITE, JCAS specific air-to-ground weapons were needed including friendly and threat indirect fire artillery, white phosphorus and colored smoke marking rounds, air-to-ground rockets, mortars,

"Katyusha" type rockets and newly deployed air-to-ground weapons like the AGM-65E Maverick laser guided air-to-ground missile. Additionally, special effects for colored non-explosive smoke markers required development. AFRL worked with the standards development communities and established protocols for smoke marking rockets and warheads to support JCAS modeling and simulation.

Most available ground target types were Soviet Era centric. More Global War on Terror (GWOT) centric targets were required. Models and scripting were developed for pickup truck mounted machine guns, civilian vehicles, single-use rocket launchers, small mortars and enemy observers.

XCITE's aircraft database was modified to allow a greater number of air-to-ground weapons loadouts. For more realistic maneuvering, an energy based aero model was added. Low altitude flight profiles and logic were added for ridge crossings. Some friendly aircraft models still require further development like AC-130 gunships, attack/observation helicopters and UAVs.

Tactical Maneuvering and Scripting

An important aspect of a CGF is its ability to accurately portray how air and ground forces move and interact with each other. Although the existing XCITE software gave instructors the ability to vector aircraft and attack ground targets, some missions required additional scripting. Aircraft on CAS missions must be able to fly to ingress and egress points, pop-up and attack ground targets and maintain restricted final attack headings. It is unreasonable to expect an instructor to control all of these behaviors, so the XCITE software was modified to autonomously fly the aircraft given mission parameters. These 3-dimensional flight profiles were significantly more difficult to script than air-to-air profiles due to the complexities of terrain interactions and dynamic maneuvering in reference to target locations. Additionally, release altitudes and dive angles for specific attacks vary greatly depending upon aircraft, weapons, terrain and tactics. As a starting point, AFRL concentrated on perfecting three generic ground attack profiles. These included a low altitude 20 degree pop-up attack, a medium attitude 30 degree dive bomb attack and a high altitude level attack replicating a precision guided bomb. AFRL engineers spent significant efforts improving scripting for these activities. Wingman flight profiles for each attack profile were also developed, but still require improvements to appear tactically realistic.

Holding and Ingress

Management of forces and airspace control are critical JTAC training tasks. Holding and attack ingress tactics were also modified to allow CGF fighters to hold at specific Contact Points (CP) points, attack from specific Initial Points (IP), attack from a right or left roll-in and return to a CP or hold at a target area. These scripts are exceptionally complex and CGF airspace management is typically still done as an IOS control input for more advanced attacks.

Coalition Scripting and Unusual Fighter Tactics

After demonstrating this attack scripting to JTAC subject matter experts, it became apparent that coalition allies employed different tactics in close air support missions than those of US pilots. For example, in actual combat British Tornado aircraft occasionally employed extremely low-altitude level attacks due to weapons and avionics requirements. Fighter and bomber aircraft are occasionally flown over target areas at low altitude and high airspeeds as a psychological show of force.

Weather Effects

A key area not fulfilled in today's DMO training environment is inclement weather effects on weapons targeting. Hot vehicle surfaces, sun angle, terrain heating and cooling, clouds and background all effect target acquisition sensors and weapon engagement zones (WEZ) of sensor targeted air-to-ground munitions. AFRL used Target Acquisition Weapon Software (TAWS), a government owned mission planning software package, to build a database of engagement zone distances for an AGM-65D Maverick missile attacking a tank from an A-10. The database was tabulated for multiple headings, altitudes, times of day, humidity, background terrain and cloud state to create a weather "Hypercube". XCITE was modified to read and check against the newly created Hypercube to obtain a validated weapons lock-on and engagement range. Although a simple demonstration on its own, it was a powerful proof of concept of how to create real-time weather affects for JCAS munitions. Before a scenario is executed, a Hypercube database of all ground targets and missile seekers could be generated under the appropriate weather conditions to support high fidelity weather based weapon engagement zones. Alternatively, the TAWS program could be stripped to a modular weather service and act as a "TAWS on demand." CGF software would request an engagement zone for any seeker against any target at any time to allow dynamic scenario changes. Work continues at AFRL to more fully develop this concept.

Database Correlation of Weapons

Although image generators have the ability to ground clamp models, munitions and detonations did not correlate perfectly. Though the IGs and XCITE constructive forces were using the exact same terrain data, how data was processed resulted in significant elevation deviations. The IG ground clamping rendered targets properly, but an air to ground missile powered by the CGF tracked to the target below the ground. On the visual system the missile fell short of the tank and detonated dozens of feet below the target. The missile properly hit the target but visually appeared as a miss. The XCITE database was switched to natively utilize the MetaVR IG's Metadesic tile data for elevations. This technique resulted in perfect correlation between the IG and the CGF models.

IOS AND SCENARIO CONTROL

To be embraced by the operational community, the instructor software had to be designed so a minimally trained JTAC could control all air and ground assets. AFRL's goal was to provide an easy to operate Instructor Operator Station (IOS) that did not require technical support for day-to-day training activities.

AFRL took the approach of implementing the JTAC's actual radio templates and call-for-fire formats into the IOS. The instructor would only have to transcribe the student's verbal control commands into the template window, select "Execute" and the mission would commence as requested. Similarly, to clear an aircraft hot or abort a mission consisted of a single click on a "Cleared Hot" or "Abort" button. Without switching between windows or navigating through menus, an instructor could model the aircraft's mission.

This first attempt at a "9-Line" JCAS briefing template worked well in demonstration, but proved insufficient for operational training. Instructors requested the ability to see more status information of the aircraft and its mission on a single screen. They specifically wanted exact time to target calculations for the scripted fighters to prevent the need to estimate the pilot's time-to-target or use manual clocks. Additional hooks were added between the IOS and XCITE to handle these on demand time-to-target calculations. By selecting the "Apply" button, the mission time would display for the instructor without commanding the aircraft. Instructors would then be able to relay to their students the first available

Figure 4. Revised JCAS 9-Line on IOS

Figure 5. CAS Override on IOS

time of attack for an aircraft. Selecting "Execute" would execute the mission and display a countdown timer as the aircraft vectored towards the target. The instructor at any time could then relay to the student over the radio the pilot's time-to-target.

During training exercises, instructors required the ability to easily change information a student had radioed without losing the student's original 9-line briefing data. A new "Override" tab was created that repeated data from the student's 9-line briefing and allowed the instructor to modify the data on the fly or to emphasize a desired learning outcome. A student could give a coordinate location of a moving target and the instructor could enter that information onto the 9-line screen. Then, as the student "talks on" the pilot, the instructor can override the called in location and select a specific entity target. The original coordinates stay recorded so during debrief the instructor can review the talk on procedure.

The override tab brings about an additional level of training for more experienced JTACs. Instructors can command the aircraft to make mistakes or react. The instructor can send the aircraft to an incorrect target, a wrong final attack heading or a different time-to-target and still save the student's original instructions. It is then up to the student to recognize the errors, compensate and abort the mission, if needed.

Laser Designation

Operationally, pilots and JTACS share laser designation information to identify targets or common reference points. In actual practice, it is difficult to hold a laser spot on a specific target due to line-of-site and pointing inaccuracies. JTACS may also designate locations near a target instead of the target itself. Simply having the entity being lased broadcast to all players that it is being designated would not fulfill all training requirements. To support these designation tactics a "laser spot" menu was devised which allows the IOS operator to lase a specific entity, a location on a database, or a small area around a point to simulate a shaking designator. The resulting DIS PDU contained information which supports the emulated GLID-II laser designator as well as simulations of other laser spot tracking systems. The laser code of the designator is also encoded in the PDU.

Artillery and Call for Fire Control

Without physics-based fly outs of artillery rounds, instructors could not properly train students to de-conflict air assets and artillery fire. Instructors needed the ability to report the time of flight of rounds and the maximum altitude the ordinance would achieve to allow the JTAC to manage artillery control airspace. AFRL continued its approach of using actual JTAC templates for the artillery call for fire missions. "Call For Fire"

and "Fire Direction Control" templates were implemented into the IOS to give instructors control of artillery assets. Similar to the 9-line, items on the list could either be typed in or selected from a drop down list. Like the initial 9-line format, this worked in a demonstration but not at an operational level.

To give instructors full control over the artillery assets, the templates were further expanded. The Fire Direction Control template was completely overhauled to allow every input given by a student on the Call For Fire tab to be modified. Figure 7 shows the target being manually edited by the instructor. Like the 9-line, the instructor can select the target the student called in on the CFF template or override with a new target location.

together. One disadvantage of the more immersive dome training system is that it permitted training only a small 2-3 JTAC team at a time. Scenario development is underway to match the existing I-FACT scenarios to the dome IOS to evaluate the training effectiveness of this system in upgrade training.

Among their criteria for scenarios, AGOS did not want the battlefield populated with static targets. Experienced JTACs quickly realized that moving targets are far more difficult for a student and the simulator could compensate for the lack of moving targets on the live range. Students would calculate a target's position but due to distractions or taskings would lose track of the enemy vehicle's location. The AGOS instructors also

Figure 6. Revised CFF on IOS

Scenario Management

The existing scenario development tools in XCITE successfully supported experienced JTACs building custom scenarios for continuation training. Scenario management for upgrading JTACs required more stringent scenario controls. The Air Ground Operations School has developed a well-defined syllabus supporting simulation training missions.

Typically, students would sit in a mass briefing where all received the same pre-briefing on that day's scenario. Using I-FACTs, six students then trained on a scenario

Figure 7. New FDC on IOS

developed scenarios that mixed high threat surface-to-air missile amongst enemy target arrays to force students to actually employ suppression of enemy air defenses fires prior to effectively conducting an airstrike.

Brief / Debrief in IOS

Debrief for air-to-air training typically involves a detailed review of the entire mission. AFRL uses DIS recorders installed on the simulation network to allow full recording of all entity actions and radio calls. After the mission the instructor can playback the entire mission or jump to a specific event. For the JCAS

Figure 8. Example Override Menu on IOS

Figure 9. JCAS Brief / Debrief System

debriefing system, AFRL utilized the same visual database and IOS as the dome to maintain familiarity. The recorder and playback utilities were built similarly to those used for typical air-to-air engagements where pilots fly for approximately one hour then debrief for one to three hours.

Observation of JTACs using the training systems found that students typically conducted a one hour mission followed by a short debrief. Additionally, instructors regularly froze the scenarios to discuss training issues as

they arose, a technique not typically used by instructors conducting air-to-air training. Since audio recordings were not made while the scenario was frozen, the debrief inevitably involved disagreements between the student and instructor as to what was said and when. The instructors were heavily tasked: controlling the scenario, acting as voice for the pilots and grading the student simultaneously. Hand written notes of student performance were written down hastily as the scenario progressed. Automated performance measurement tools and immediate feedback may be more useful in future systems than full-scenario playback capabilities, though full-scenario playback should still be available for more complex DMO events.

Figure 10. JTACS Training in Immersive Dome

AFRL is working to introduce automated real-time DIS speech to text transcription of the scenarios. The instructors could then refer to the transcript for a no-argument "you said this" during debrief with the students. Students would be able take their transcripts with them when they leave so they can further review what they did right and wrong in the mission. Additionally, a secondary radio frequency could be setup for the instructor to allow him to make comments as the mission progressed that the student would be unable to hear. After the mission those comments could be played back or read from the transcript.

Scenario Generation for ROVER Training

The requirement for training indirect control of JCAS assets was highlighted in previous sections. The United States Air Forces in Europe Warrior Preparation Center developed a method that allows unique training with the ROVER system. A predator UAV was flown using the Air Force Synthetic Environment for Reconnaissance and Surveillance / Multiple Unified Simulation Environment (AFSERS/MUSE) which supported a sensor representation through a network connection to a ROVER laptop computer. XCITE was used to generate

targets, strike aircraft and munitions. Correlation between the ROVER sensor visualization and the XCITE CGF was excellent. This system has provided superb training to develop advanced tactics and prepare for combat deployments and demonstrates the potential for interfacing multiple CGFs to provide targeted training activities for advanced systems.

LIVE-VIRTUAL-CONTRUCTIVE JCAS

In 2007, AFRL showcased a Live Virtual Constructive (LVC) demonstration at the Air Force Association and Interservice/Industry Training, Simulation, & Education conferences. In these demonstrations, a transportable 5 meter JTAC dome along with two deployable F-16 cockpits were setup on the exhibit floor. Utilizing ACMI pods and Link-16 connections, the JTACs within the dome were able to see and control the live aircraft flying throughout the DMO environment. The JTACs real radio was linked with emulation software to transmit the data over the DIS network and the live F-16 pilots used their UHF radios to transmit to a similar conversion device at Luke AFB.

Figure 11. Live Aircraft at Luke AFB

Although the interactions between the pilot and JTACs were real, the interactions with the range targets were not. Ground targets in the DMO environment could easily be engaged any time using the XCITE software, but those entities would not appear on the live range or on the instrumentation inside the F-16. A Link16 connection did permit XCITE air assets to appear on the datalink displays in the live aircraft.

Even though the F-16s were dropping real munitions at the range, weapons release data could not be passed to the JTAC Dome over unclassified lines. To allow the JTAC to observe weapons effects, a "magic bomb" was added to the IOS which allowed the instructor to drop a

bomb at any location at any time within the simulation. A classified LVC connection would have permitted information such as weapon release to be relayed over the simulation network. In this case, the CGF could be switched to a weapons server to display a simulated weapons flyout over the network. It should be noted that any small errors due to latency, data dropouts or maneuvering would cause huge differences between
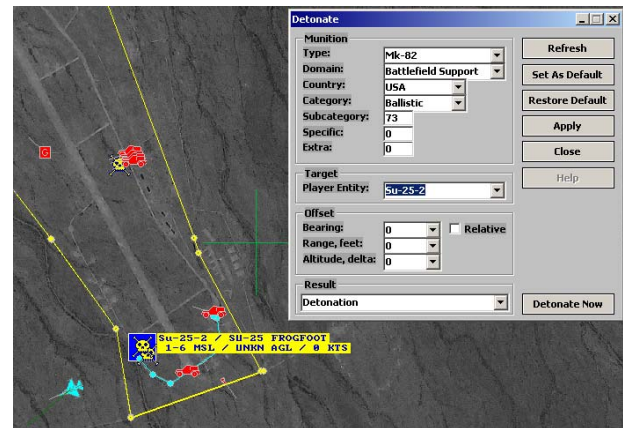
Figure 12. Magic Bomb on IOS

where the bomb actually dropped and where the simulation calculated its drop. One potential solution under consideration is to have scoring plots of actual bomb impacts mapped into the LVC network to display a correlated bomb impact. Further work is required in this area.

FUTURE REQUIREMENTS

AFRL has identified current technical shortfalls relating to JCAS training systems. The existing training system can provide only limited interactions with actual ground command and control agencies. Most interactions, like artillery fire support, are controlled by a role playing JTAC. In the future, improved command and control modeling, night and adverse weather representations, models for advanced weapons and weapons effects and seamless integration with existing CGFs in high entity count scenarios are required.

Integration to Joint Fire DMO Environments

AFRL's CGF development centered on providing models and simulations specific to Air Force JCAS Training Research. Integration with actual US Army constructive simulations and training systems is desired to fully represent the entire Theater Air Ground System. Interfaces to validated Army and Special Operations models and simulations should be developed to employ

a "best of breed" approach for constructive forces support. An optimal mix of constructive forces would use air centric CGFs for aircraft, air delivered munitions and enemy surface-to-air threats while using ground centric CGFs for vehicles, convoy routing, artillery weapons and ground command and control like Blue Force Tracking, Fire Support Cells and tactical ground force command and control. Rapid integration and correlation between systems is desired.

Automated Command and Control for Rapid Scenario Generation

In high entity count scenarios, technologies that automate scenario generation, manage ground force-on-force activities and provide synthetic C2 are desirable. The Theater Air Ground System Synthetic Battlespace is an example of efforts to automate scenario generation and provide theater level of war command and control support to live virtual constructive training systems (Ales, 2006).

Improved Nighttime Simulation

The JTAC TRS system developed by AFRL did not display high fidelity, validated night vision scenes. Future JTAC training systems will require night vision representations. In this case, CGFs must be modified for both ground and air models to provide night tactics and target representations. This would include lights-on and lights-off convoy movements, modeling of target acquisition ranges for night vision and additional infrared sensors, night formation tactics for aircraft and support for night visual special effects like tracer fire. Models to support artillery and air delivered parachute flares and markers are also required.

Damage States for Models and Munitions Effects

In current operations, urban CAS and operations in cluttered terrain are the norm. A training requirement exists to mange firepower and prevent collateral damage and fratricide in urban JCAS. Due to the destructive force of air delivered munitions, precise modeling of damage effects to buildings and other representations of collateral damage could provide useful training feedback. Warhead effects need to be modeled extremely accurately and validated for precision engagement in urban terrain.

CONCLUSION

AFRL successfully demonstrated modification of an air centric constructive training environment to support a high fidelity joint close air support training system. Future acquisitions for JCAS training systems should study AFRL's lessons learned and ensure realistic models, scripting, air-to-ground tactics and realistic artillery control are available. Capabilities to support growth in advanced and coalition tactics must also be considered. Instructor operating requirements for JCAS vary greatly from those of aircraft simulators and combining scenario control features for both air and ground models in a single system is desirable. Involving constant feedback from JCAS subject matter experts while developing computer generated forces and instruction operating systems is possibly the most critical step to ensuring usability and requirements goals.

REFERENCES

Ales, Ricky R. & Steven G. Buhrow. (2006). *The Theater Air Ground System Synthetic Battlespace*, Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2006.

Erwin, Sandra I. (2008), Tough Calls, *National Defense, May, 2008*, 46-49.

Kauchak , Marty (2008), A Significant Improvement in the Way They Train, *Military Training Technology*, Volume: 13 Issue: 2, Apr 13, 2008.

United States Air Force (2007), *Tactical Air Control party Enabling Concept.*

United States Department of Defense (2007), *Joint Close Air Support (JCAS) Action Plan.*

United States Government Accounting Office (2003), GAO-03-505, *Lingering Training and Equipment Issues Hamper Air Support of Ground Forces.*
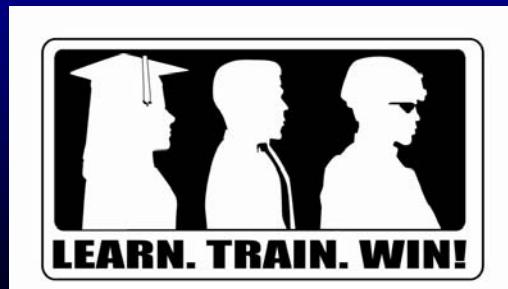
This page intentionally left blank.

I/ITSEC 2008
Computer Generated Forces for Joint Close Air Support and Live Virtual Constructive Training

Craig Eidman, AFRL and Lt Clinton Kam, AFRL
craig.eidman@mesa.afmc.af.mil
clinton.kam@mesa.afmc.af.mil

Topics

- **Joint Close Air Support (JCAS) training requirements and shortfalls**

- **Training research program history**

- **Computer Generated Forces (CGF) shortfalls and improvements**

- **Scenario control**

- **Live-Virtual-Constructive (LVC) interactions**

- **Lessons learned and future requirements**

JCAS Training Overview

- **Live range shortfalls**
  - **Static and non-JCAS targets**
  - **Training munitions / dry passes**
  - **Safety / range restrictions**
  - **Scheduling / cost**

- **Simulation advantages**
  - **Greater frequency, lower cost**
  - **Focused training**
  - **Advanced training opportunities**

The live fire range is often a limited representation of actual joint combat ops … simulation is a reasonable alternative

Training Research Program

- **Part-task JCAS training solutions**
  - **Focused solely on Joint Terminal Attack Controller (JTAC)**
  - **Low fidelity models**
  - **Limited FOV and distributed training**

- **Fully-immersive JCAS training solutions**
  - **360x180 field of view dome**
  - **Actual radios & GPS equipment**
  - **Simulated sensor devices**
  - **Full Distributed Mission Operations (DMO) DIS network connectivity**

CGF Shortfalls & Modifications

- **Marking and artillery**

- **Air-to-ground weapons modeling**

- **Global War on Terror (GWOT) relevant threat models**

- **Air-to-ground tactics, maneuvering, and scripting**

- **Coalition forces models**

- **Weather / environmental effects**

- **Database and detonation correlation**
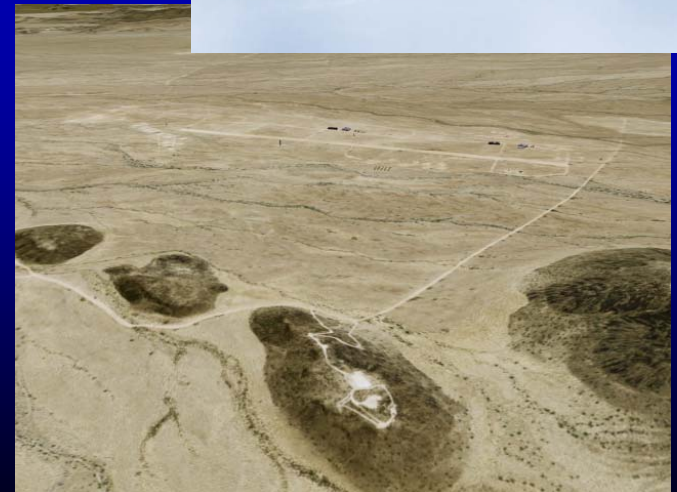
An attack example ...
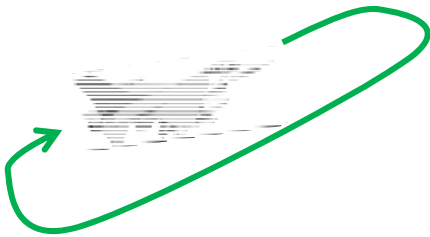
Holding at Initial Point (IP)

An attack example ...
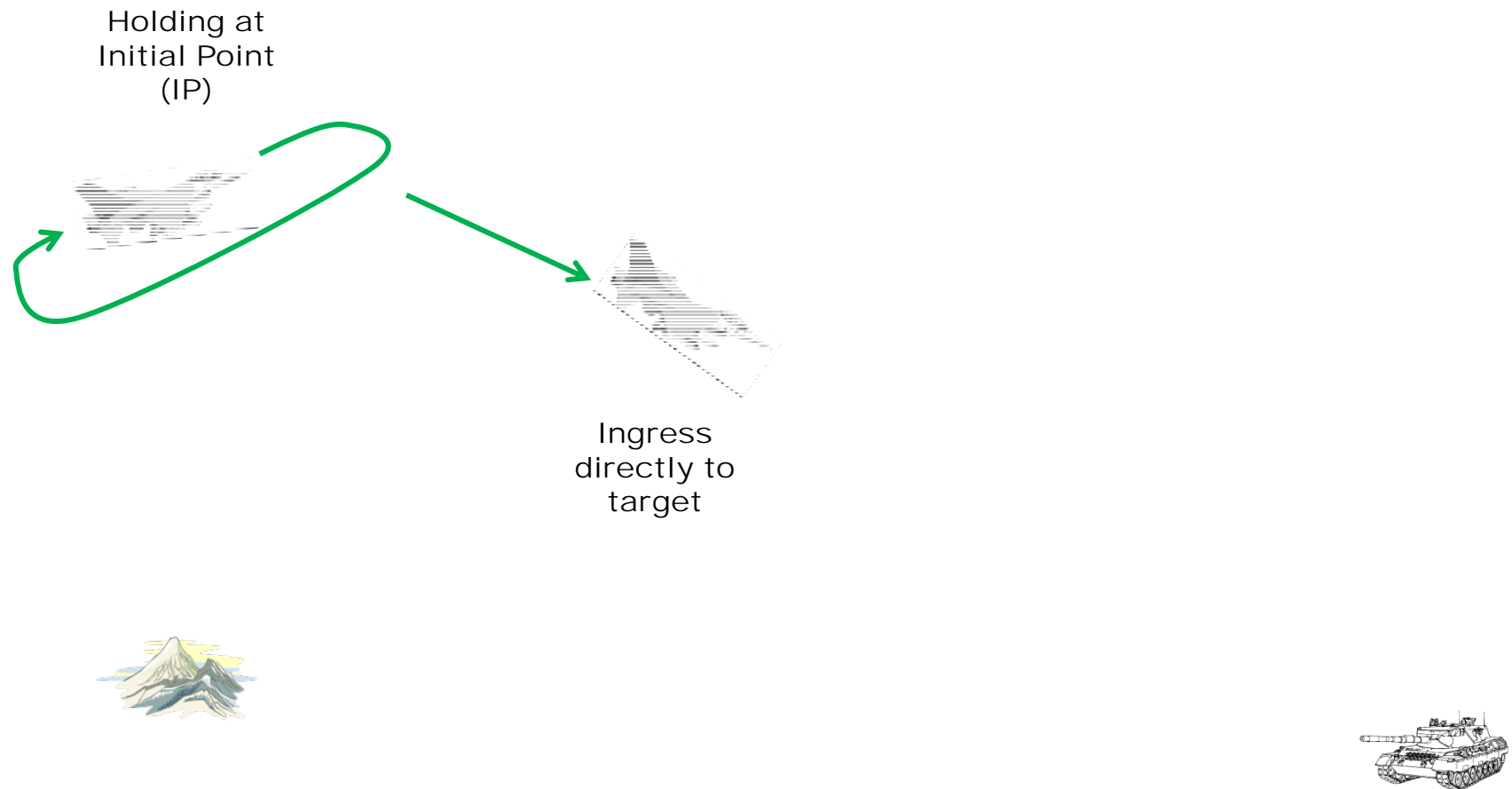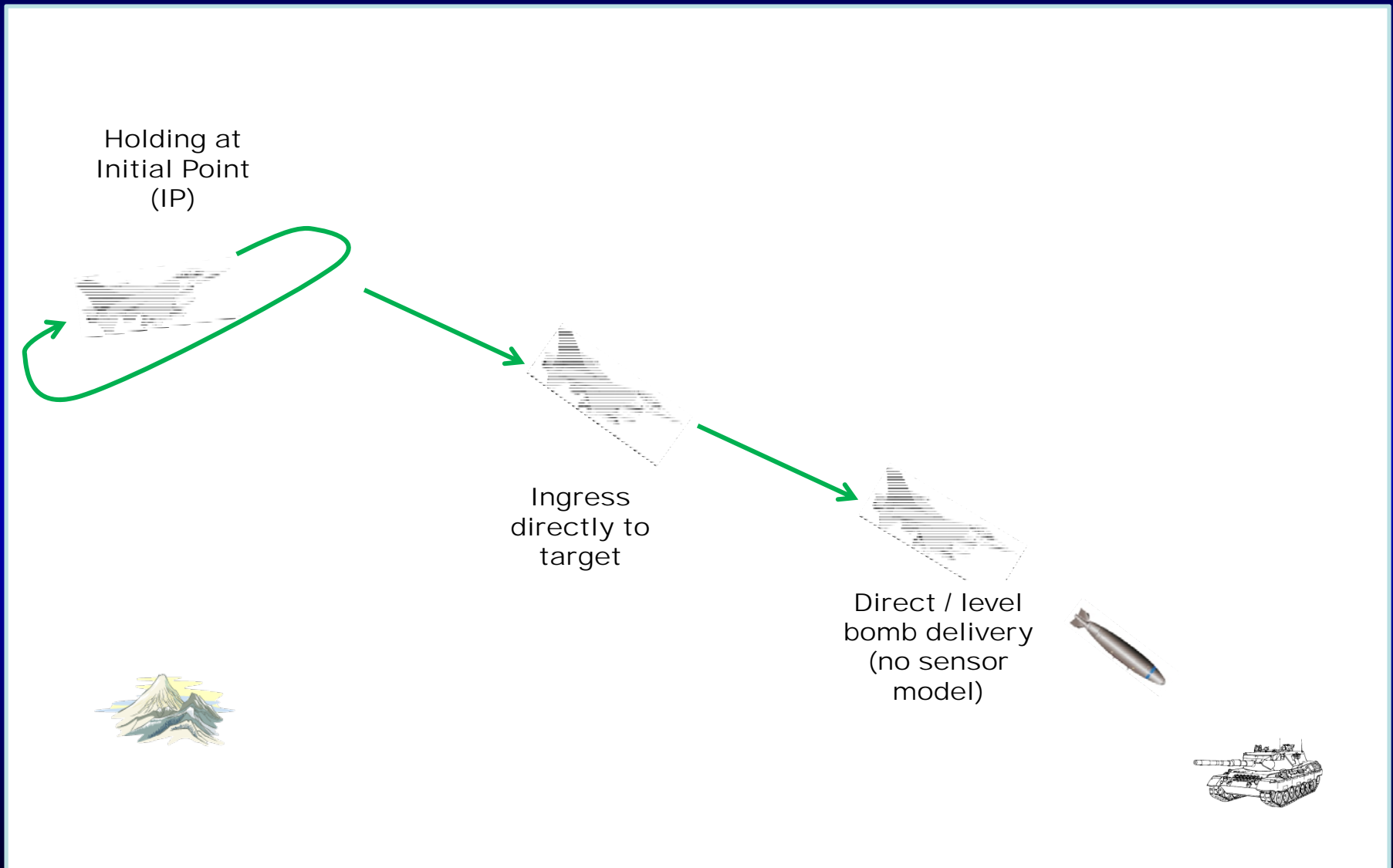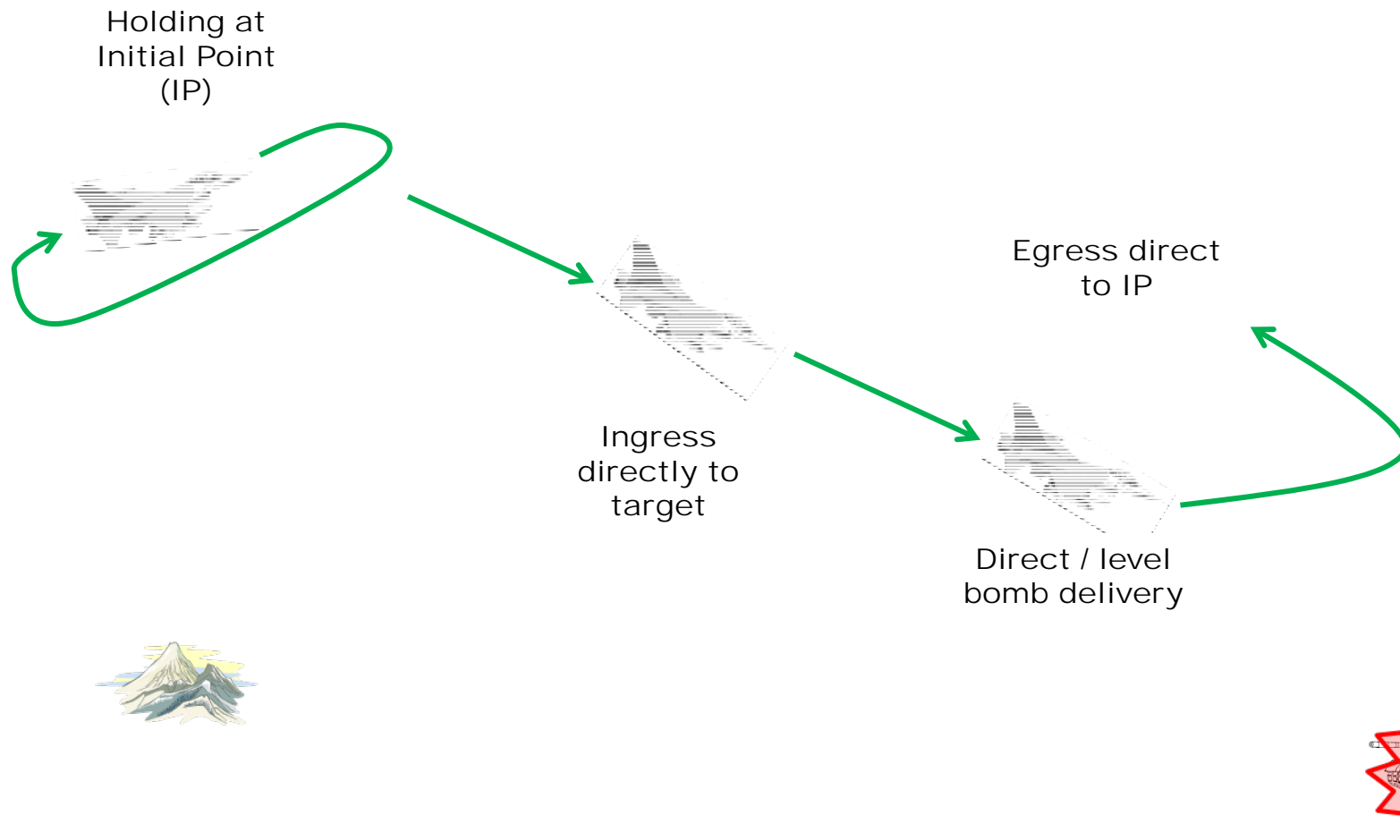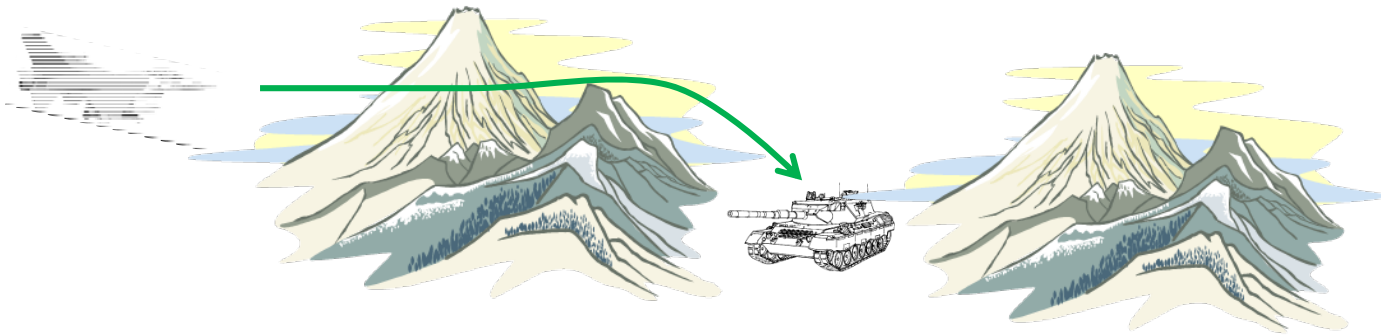
Holding at
Initial Point
(IP)

Ingress
directly to
target

An attack example …

Holding at
Initial Point
(IP)

Ingress
directly to
target

Direct / level
bomb delivery
(no sensor
model)

An attack example ...

Holding at
Initial Point
(IP)

Egress direct
to IP

Ingress
directly to
target

Direct / level
bomb delivery

Altitude Problems
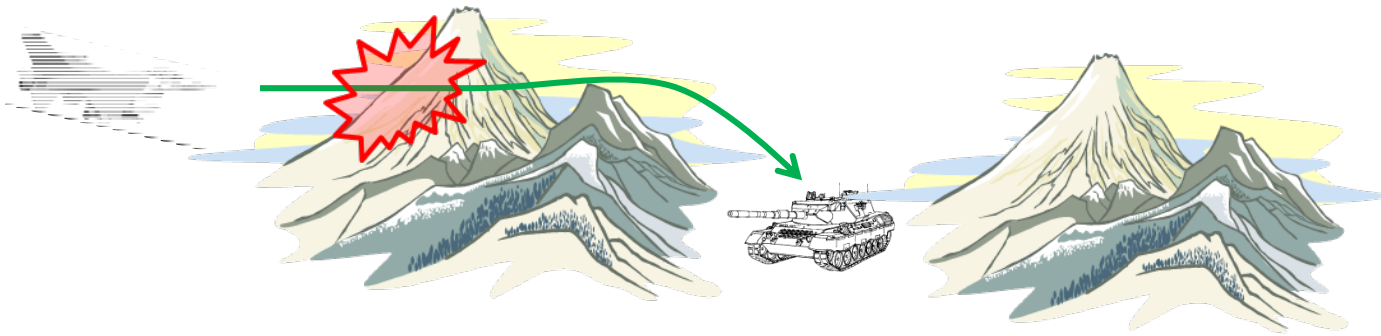
Altitude Problems
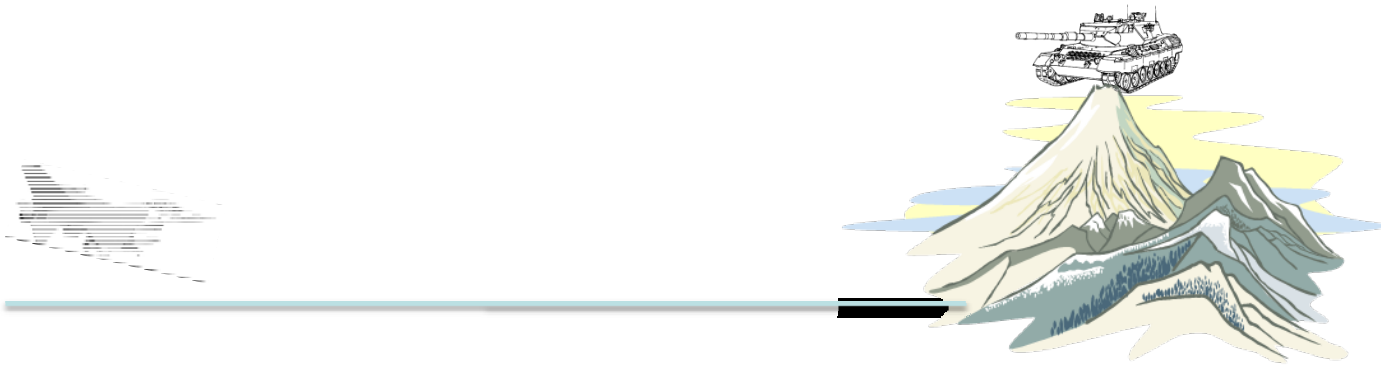
Altitude Problems
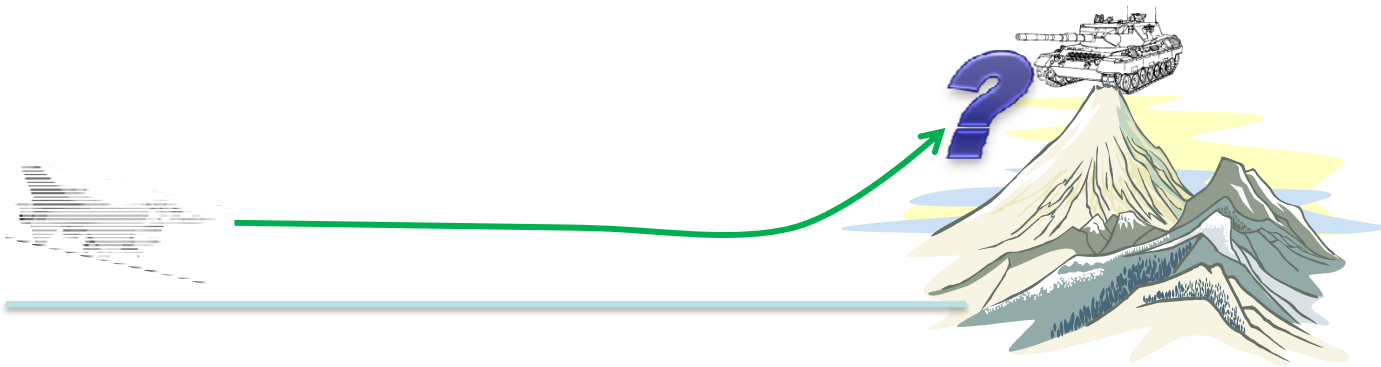
Altitude Problems
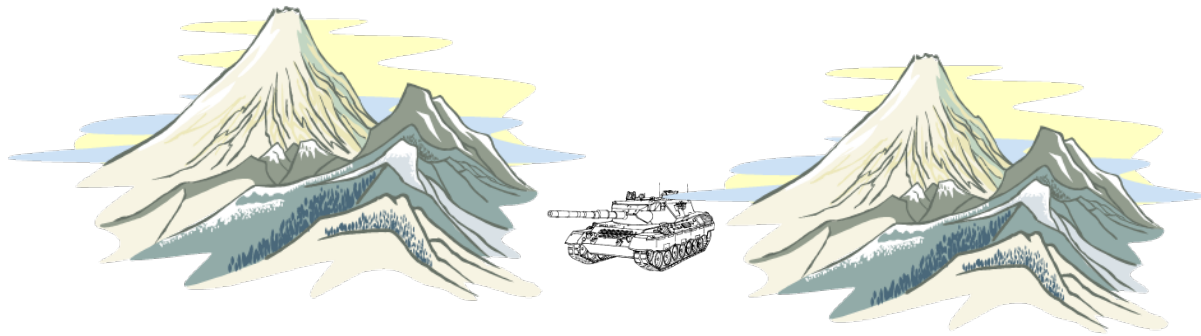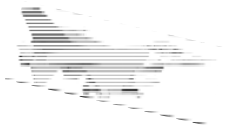
Altitude Problems

Altitude Problems
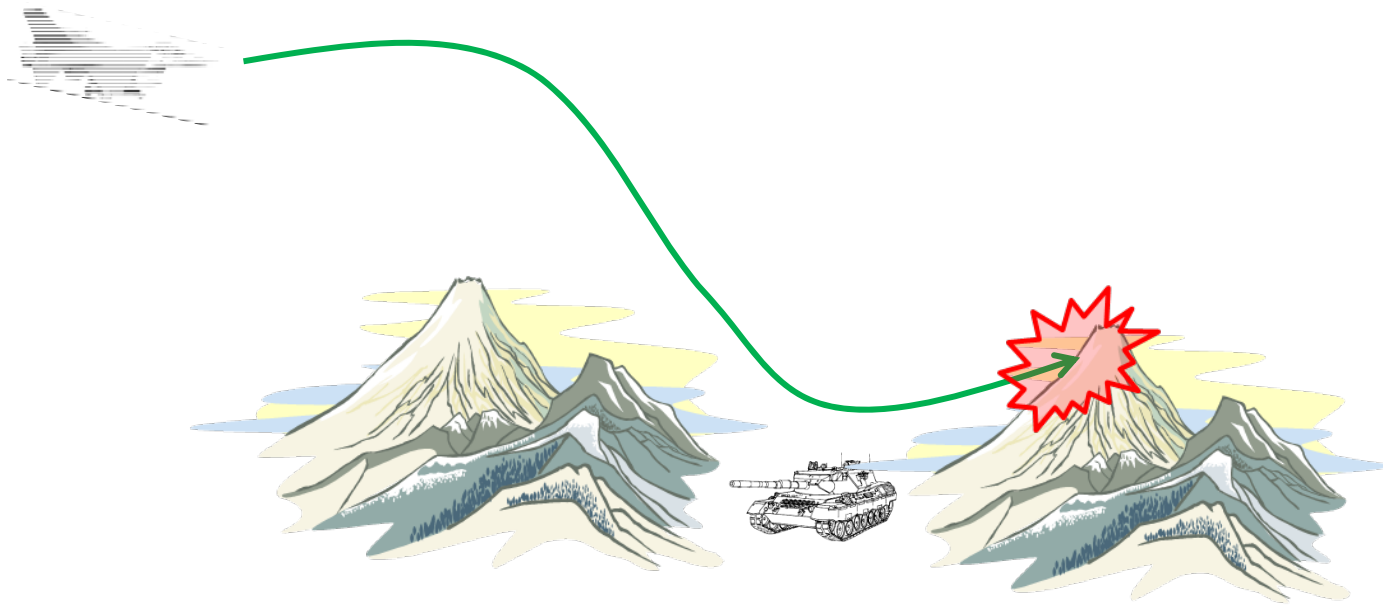
Improved A-G Scripting

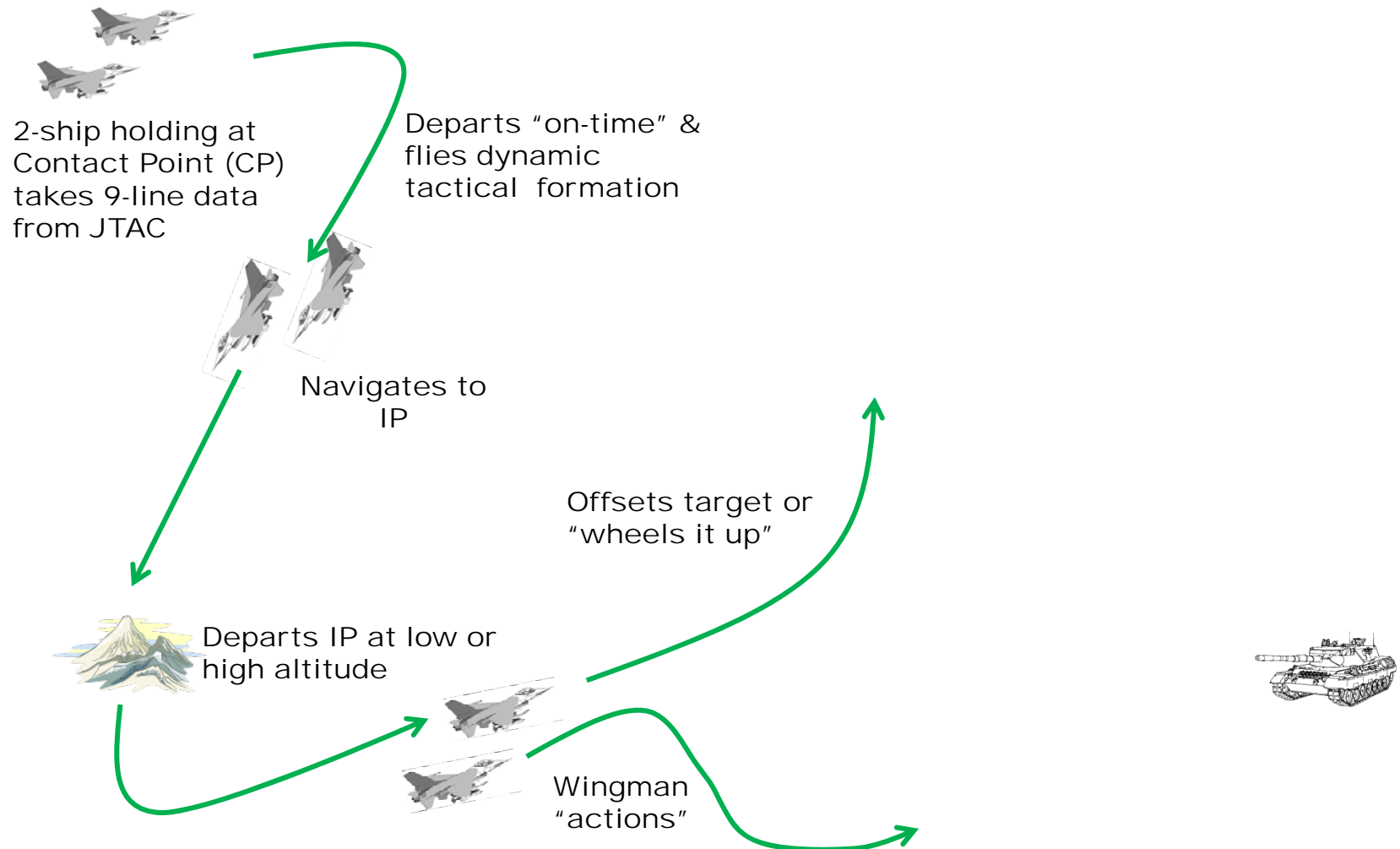**2-ship holding at Contact Point (CP) takes 9-line data from JTAC**

Improved A-G Scripting

2-ship holding at Contact Point (CP) takes 9-line data from JTAC

Departs "on-time" & flies dynamic tactical formation

Navigates to IP

Improved A-G Scripting

2-ship holding at Contact Point (CP) takes 9-line data from JTAC

Departs "on-time" & flies dynamic tactical formation

Navigates to IP

Departs IP at low or high altitude

Improved A-G Scripting

2-ship holding at Contact Point (CP) takes 9-line data from JTAC

Departs "on-time" & flies dynamic tactical formation

Navigates to IP

Offsets target or "wheels it up"

Departs IP at low or high altitude

Wingman "actions"

Improved A-G Scripting

2-ship holding at Contact Point (CP) takes 9-line data from JTAC

Departs "on-time" & flies dynamic tactical formation

Navigates to IP

Roll-in, sensor or target aq model

Departs IP at low or high altitude

Offsets target or "wheels it up"

Intercepts dive bomb path, strafe, or Maverick

Wingman "actions"

Improved A-G Scripting

2-ship holding at Contact Point (CP) takes 9-line data from JTAC

Departs "on-time" & flies dynamic tactical formation

Navigates to IP

Roll-in, sensor or target aq model

Egress as directed

Offsets target or "wheels it up"

Intercepts dive bomb path, strafe, or Maverick

Departs IP at low or high altitude

Wingman "actions"

Wingman attacks

Scenario Control

- **9-Line templates**
  - **Status info, TOT, override**

- **Laser designation tactics**

- **Artillery and call-for-fire**

- **Scenario management**
  - **Vignette time, moving tgts**

- **Brief / debrief**

- **ROVER training**

Live-Virtual-Constructive

- **2007 AFA and I/ITSEC demonstrations**

- **ACMI/Link-16 connections**

- **Radio communication**

- **Weapons release / Magic Bomb**
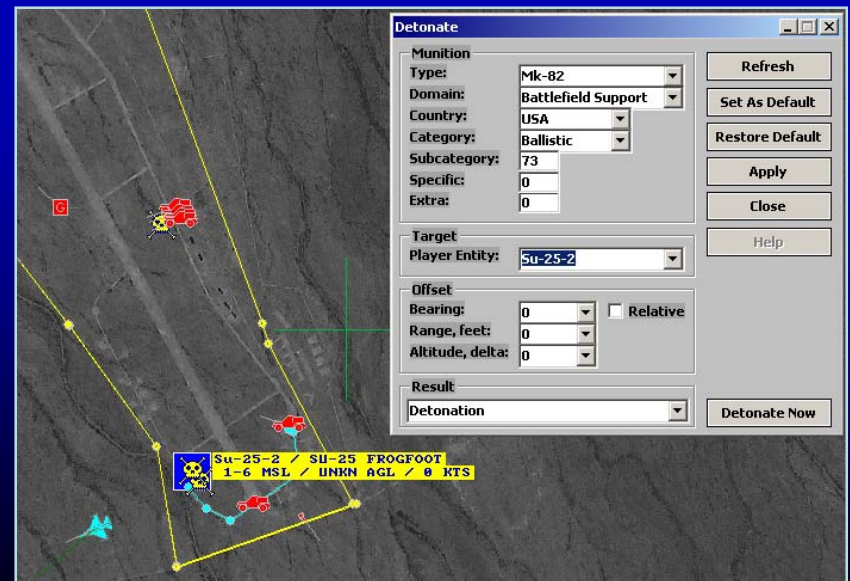
JCAS Training Lessons Learned

- **Air-to-ground JCAS modeling and scripting is significantly more challenging than air-to-air**
  - More "fly by the seat of the pants"
  - More 3-dimensional
  - Must always reference terrain, target, weapons parameters
  - Requires near perfect database correlation
- **Attacks must look realistic for valid training**
- **Short scenarios with "lessons learned" discussed between events**
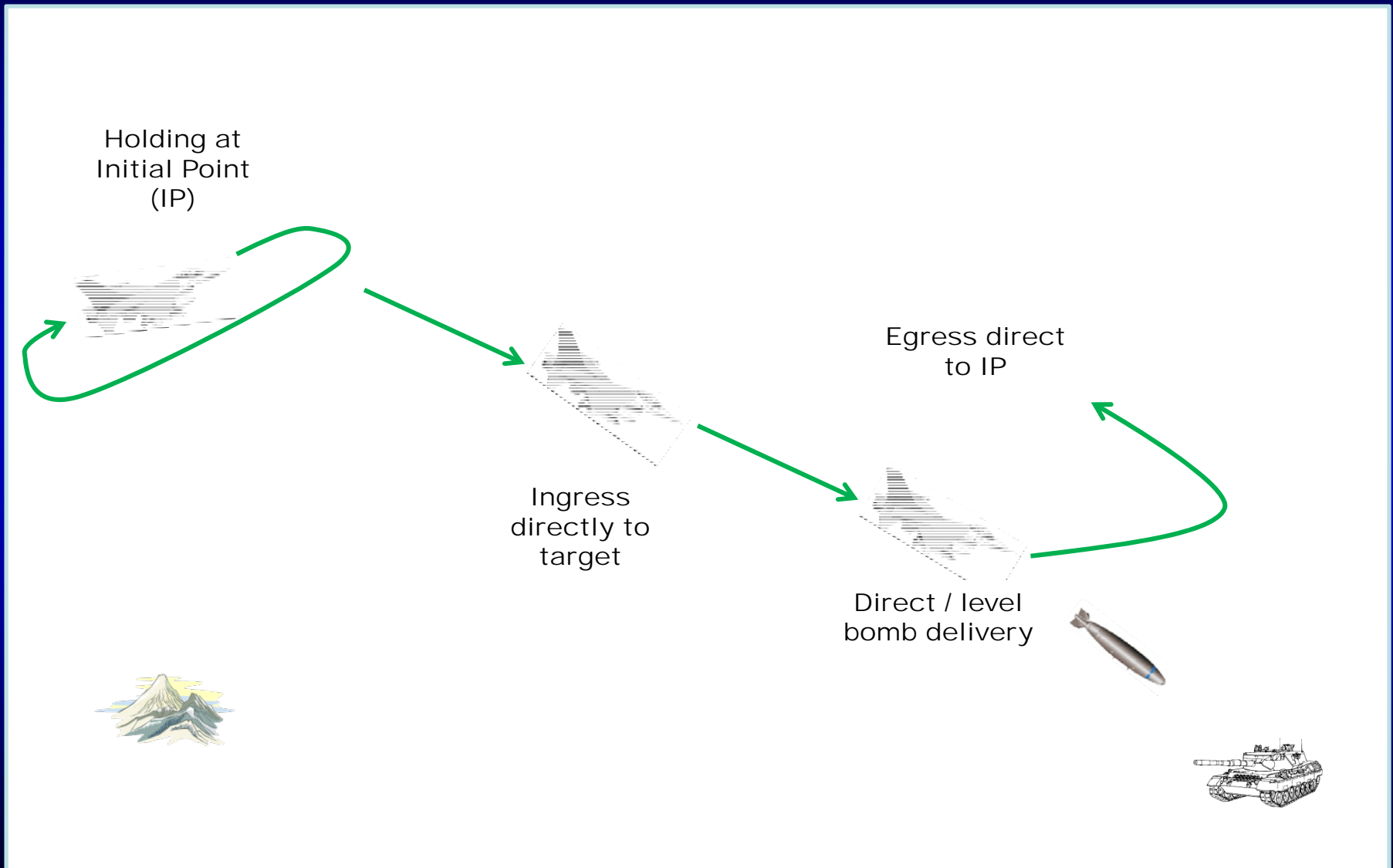- **IOS feeds data to instructors at real time**

Future Requirements

- **Continued scripting improvements, tactical models, and AI**

- **Integration of Joint Fire DMO events**

- **Automated C2 for high entity counts**

- **Improved nighttime simulation**

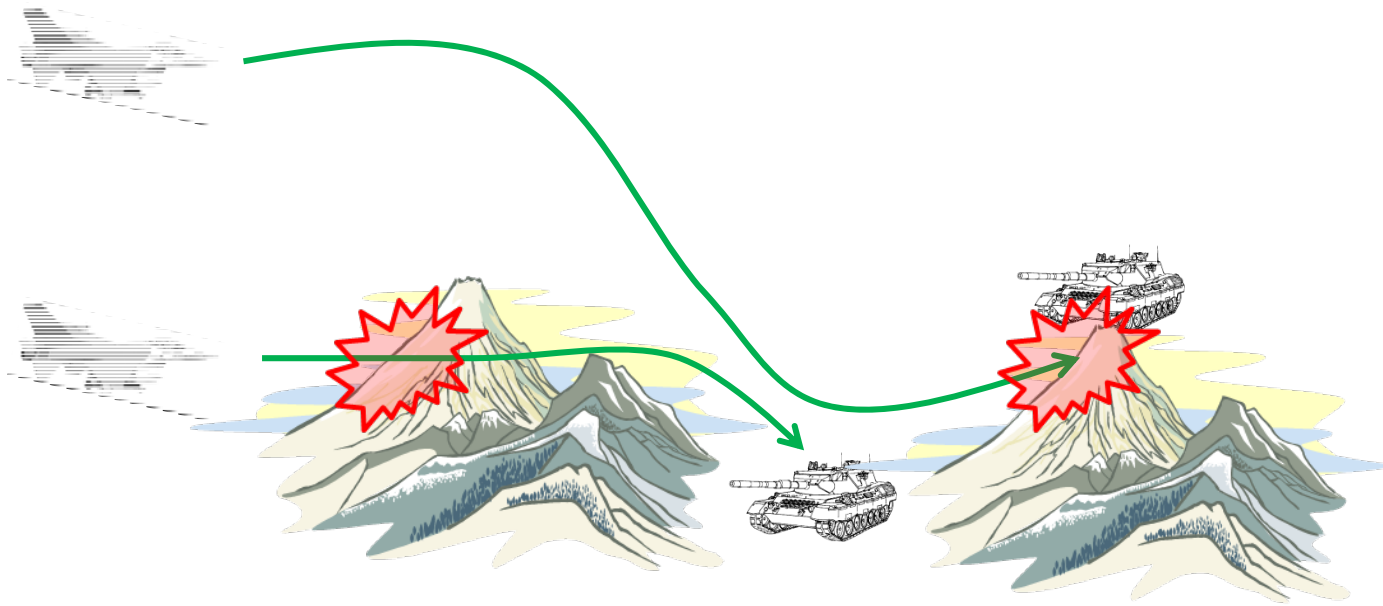- **Damage states for models / munitions effects**

Questions?

- **JCAS training requirements and shortfalls**

- **Training research program history**

- **Computer Generated Forces improvements**

- **Scenario control**

- **Live-Virtual-Constructive interactions**

- **Lessons learned and future requirements**

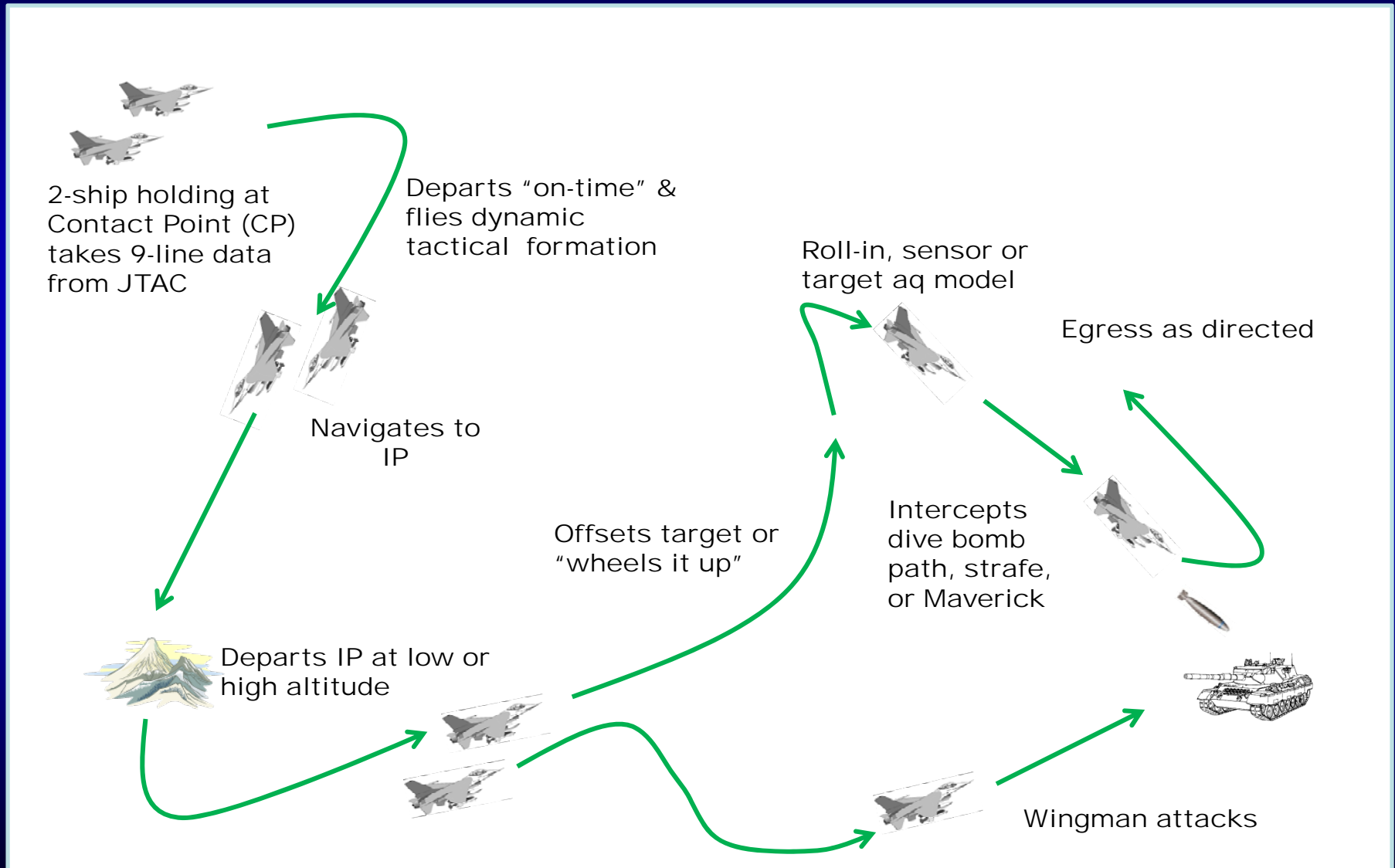- *Demonstrations available in the US Air Force (Quad) booth #1923*

An attack example ...

**Holding at
Initial Point
(IP)**

**Egress direct
to IP**

**Ingress
directly to
target**

**Direct / level
bomb delivery**

Altitude Problems

Improved A-G Scripting

2-ship holding at Contact Point (CP) takes 9-line data from JTAC

Departs "on-time" & flies dynamic tactical formation

Navigates to IP

Departs IP at low or high altitude

Offsets target or "wheels it up"

Roll-in, sensor or target aq model

Intercepts dive bomb path, strafe, or Maverick

Egress as directed

Wingman attacks

This page intentionally left blank.

Assessing High-Fidelity Training Capabilities Using Subjective and Objective Tools

Leah J. Rowe
L-3 Communications
Mesa, AZ
leah.rowe@mesa.afmc.af.mil

Justin H. Prost
Lumir Research Institute
Mesa, AZ
justin.prost@lumir.com

Brian T. Schreiber
Lumir Research Institute
Mesa, AZ
brian.schreiber@lumir.com

Winston Bennett, Jr.
Air Force Research Laboratory
Mesa, AZ
winston.bennett@mesa.afmc.af.mil

ABSTRACT

Instructors often assess training effectiveness using subjective evaluation tools. The use of evaluation by Subject Matter Experts (SMEs) assumes that the experts can distinguish between small but meaningful differences in the measured domain. Subjective evaluations by experts provide both an efficient and effective means of identifying the strengths and weaknesses of the assessed entity. In the area of simulation development, SME assessments evaluate the training capabilities of systems, identify deficiencies, and compare the relative impact of the various deficiencies. This paper presents methods that utilize subjective assessments from SMEs and compares SME ratings of Mission Essential Competency (MEC) experiences with objective performance measures. The methodology entails mapping the correspondence between MECs and objective performance measures. Additionally, we mapped performance measures to training scenarios in order to determine the appropriate skills for evaluation. This study uses performance measures based on the capabilities of the simulators in our laboratory. The congruence of the subjective evaluations by experts and objective simulator performance variables provides validation for the use of subjective assessments completed by experts. The results provide a strong framework for building an understanding of the relationship between subjective and objective performance data to measure training effectiveness.

ABOUT THE AUTHORS

Leah J. Rowe is a Research Scientist with L-3 Communications at the Air Force Research Laboratory, 711th Human Performance Wing in Mesa, AZ. She completed her M.S. in Applied Psychology at Arizona State University in 2007. Leah is presently pursuing a Ph.D. in Industrial/Organizational Psychology at Capella University.

Justin H. Prost is a Research Scientist with Lumir Research Institute. He completed his Ph.D. in Developmental Psychology at Arizona State University in 2001. Recently, he has worked on current simulation research at the Air Force Research Laboratory.

Brian T. Schreiber is CEO and Senior Scientist with Lumir Research Institute in support of the Air Force Research Laboratory, 711th Human Performance Wing, in Mesa, AZ. He completed his M.S. in Human Factors Engineering at the University of Illinois at Champaign-Urbana in 1995.

Winston Bennett, Jr. is a Senior Research Psychologist and team leader for the training systems technology and performance assessment at the Air Force Research Laboratory, 711th Human Performance Wing, in, Mesa AZ. He received his Ph.D. in Industrial/Organizational Psychology from Texas A&M University in 1995.

Assessing High-Fidelity Training Capabilities Using Subjective and Objective Tools

Leah J. Rowe
L-3 Communications
Mesa, AZ
leah.rowe@mesa.afmc.af.mil

Justin H. Prost
Lumir Research Institute
Mesa, AZ
justin.prost@lumirresearch.com

Brian T. Schreiber
Lumir Research Institute
Mesa, AZ
brian.schreiber@lumirresearch.com

Winston Bennett, Jr.
Air Force Research Laboratory
Mesa, AZ
Winston.bennett@mesa.afmc.af.mil

INTRODUCTION

Assessment systems, training programs, and subjective assessment tools are the product of expertise. To become an expert, one must obtain both skills and knowledge in a specific domain (Schvaneveldt, Tucker, Castillo, & Bennett 2001). We rely on subject matter experts (SMEs) in many fields (e.g., law enforcement, human factors, medicine, and engineering). The military is no exception to this rule, and uses SMEs regularly.

SMEs have knowledge, skills, and experiences that set them apart from the average field practitioner. They can identify subtle cues that less-experienced operators may miss during complex tasks and in specific environments. SMEs often provide simple assessment solutions for very complex measurement tasks (Schreiber, Gehr, & Bennett, 2006).

Yet even a SME, may find it difficult to assess performance effectively. Historically, Warfighter performance has been assessed using subjective grading measures either by SMEs or Instructor Pilots (Schreiber, et al., 2006; Krusmark, Schreiber, & Bennett, 2004; Crane, Robbins, & Bennett, 2000). Researchers continually strive to identify or create objective performance measures. At the Air Force Research Laboratory (AFRL) in Mesa, Arizona, researchers have developed a system that collects objective data from a complex high-fidelity simulation environment. This paper discusses a method of combining objective and subjective data to assess training research in the Distributed Mission Operations (DMO) Training Research Testbed (TRT) at AFRL Mesa.

We begin by discussing the differences between subjective and objective data, and highlight the advantages of each. Next, we discuss the AFRL DMO TRT highlighting the approach that combines subjective and objective data to create a metric to measure training effectiveness. Finally, we discuss the methodology used, findings, and implications for the future.

Subjective versus Objective Performance Assessment

Subjective Data

Subjective data provides the only means for assessing both opinions and preferences. Subjective data is collected frequently as it is typically easy to obtain and inexpensive, these two factors may influence practitioners when they select a data collection method (Cushman & Rosenbery, 1991). Nevertheless, in some situations subjective data is the only data source that is available or feasible.

At the DMO TRT, we collect both subjective and objective performance data. F-16 SMEs generate the subjective data by completing SPOTLITE (Scenario-based Performance Observation Tool for Learning in Team Environments). SPOTLITE allows observers to measure and assess team and individual performance in live and simulated training exercises in real time (MacMillan, Entin, Morely, & Bennett, under review).

Objective Data

Researchers often prefer objective data in research, because it ideally lacks bias; however, it is often difficult to obtain. To be truly objective, there must be an "absolute" answer absent of human opinion. This situation in itself creates a barrier when building objective assessments. In addition, objective measures are generally more costly and time consuming than subjective measures (Cushman & Rosenbery, 1991).

In the DMO TRT, we collect objective performance data with the Performance Evaluation Tracking System (PETS). PETS provides the Warfighter with exact data regarding their actions during live and training events

by collecting and distilling millions of data points directly from the simulator (Schreiber & Bennett, 2006). We describe PETS in more detail below.

Which Assessment Method to Use?
PETS gathers micro-data that is not feasible for a human to track, whereas SPOTLITE assesses performance with criteria that only a SME can assess. It is necessary to identify the most appropriate assessment method for any performance evaluation. The fundamental differences between PETS and Spotlite make it clear that performance assessment does not fall in a "one size fits all" category.

Subjective assessments often prove to be the most efficient mechanism for obtaining information; however, when subjective assessments are appropriate, it is important to assure data quality by gathering it from a reliable source. SMEs have expertise that improves the reliability of subjective data

In prior research, objective data showed that, F-16 pilot performance improved from pre- to post-training in the DMO TRT (Schreiber & Bennett, 2006; Rowe, Gehr, Cooke, & Bennett, 2007). Additionally, subjective measures showed that pilot knowledge changed from pre- to post-training in the DMO TRT as well (Rowe, Gehr, Cooke, & Bennett, 2007; Rowe, Schvaneveldt, & Bennett, 2007).

This paper presents an approach to mapping subjective F-16 SME ratings to objective performance data. Building a process that integrates SME evaluations and objective performance data will allow integration of more sophisticated training protocols in the DMO environment. In any training environment, SMEs are limited to what they can observe. The DMO TRT has more performance information available, a result of both technological advances (e.g. objective performance measurement tools) and the increased number of participants. Providing instructors with objective performance measures will allow development of more effective and efficient training protocols. One such example is the development of "adaptive training."

Distributed Mission Operations Training Research Testbed

DMO Defined
DMO is a system of networked simulators that supports multi-player training for combat exercises. DMO is different from stand-alone simulation systems, such as those used to train emergency procedures, in that it provides combat-like experiences involving real-time interaction with other entities, both virtual (e.g., a flight wingman in another simulator) and constructive (e.g., hostile entities). The objective of DMO is to train higher-order skills and improve team coordination while executing significant portions of an entire mission (Colegrove & Alliger, 2002).

The DMO TRT consists of four high-fidelity F-16 simulators, a high fidelity Air Battle Manager Simulator, a computer-generated threat system, and an instructor/operator station. The DMO TRT also includes a well equipped brief/debrief room (the DMO TRT is shown in Figure 1).

Figure 1. Overall view of Mesa AFRL DMO Training Research Testbed

Mission Essential Competencies
Syllabi trained in the DMO TRT are structured based on Mission Essential Competencies (MECs), defined as "higher-order individual, team, and inter-team competencies that a fully prepared pilot, crew or flight requires for successful mission completion under adverse conditions and in a non-permissive environment" (Colegrove & Alliger, 2002, p. 12). A competency-based training structure defines a standard level of proficiency or competency that one must have in order to be efficient in his/her job, thus emphasizing ways to address deficiencies in skills, knowledge, or experience in individuals, teams, or crews (Schreiber & Bennett, 2006).

Performance Evaluation Tracking System
PETS developed at AFRL, as an Advanced Technology Demonstration for the Air Combat Command, is a software tool that enables multi-platform, multi-level measurement at the individual, team, and inter-team levels in complex, live, virtual, and constructive environments (Schreiber & Bennett, 2006).

Installed in the DMO TRT PETS collects, stores, and organizes up to one million data points per minute. Schreiber and Bennett (2006) validated the use of PETS in a simulated environment. Additionally, they were able to define the most sensitive air-to-air measures for the F-16 in this environment, meaning the measures that are most significantly impacted from pre- to post training in the DMO TRT.

METHODS

Participants

Two hundred-seventy-two F-16 fully qualified F-16 pilots from United States Air Force, Air National Guard, and Air Force Reserve pilots participated in this study. The pilots consisted of 53 teams or four or five pilots each. Their mean age was 33.1, and they had an average of 10.8 years of military service and 1,016 F-16 flight hours.

Another sample consisted of seven F-16 SMEs. All participants were male, with a mean age of 40.8 years. Two are active in the Air National Guard and five retired from the Air Force between one and two years ago.

Procedures

DMO Training Research Week
Each team participated in nine 3½-hour training sessions over the course of the single DMO training week. Each session included a one-hour briefing, an hour of flying multiple engagements of the same mission genre, and a 90-minute post-mission debrief. Syllabus scenarios were either offensive or defensive, and consisted of four F-16s versus a varying number of threats. The team flew three benchmark scenarios at the beginning of the week and again at the end of the week for evaluation purposes.

Flight Performance
We assessed flight performance using PETS. Metrics were derived to measure performance change in three areas: weapons employment, weapons engagement zone management, and overall performance.

The benchmarks were constructed as scenarios where the four-ship of F-16s and their Air Battle Manager defended against eight threats (six hostiles and two strikers). All benchmarks were designed to be of equal complexity. We randomly assigned each team three-benchmark scenarios. The participants flew in the same cockpits during all benchmark scenarios. On day five, teams flew mirror image missions of the three benchmarks. Figure 2 illustrates a benchmark and its

mirror image. All of the benchmark scenarios that were utilized during this research are equally complex (Denning, Bennett, & Crane, 2002).
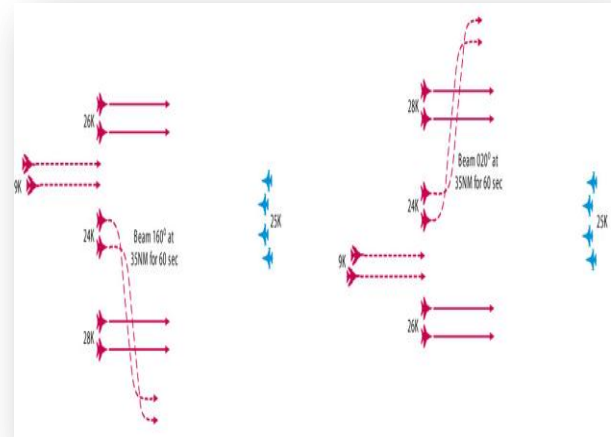

Figure 2. Mirror-Image Point Defense Benchmark Scenarios

Knowledge, Skill, and Performance Mappings
F-16 SMEs completed three sets of ratings to complete the tasks described in the following paragraphs. Each task utilized an identical Likert scale (0 = Not Relevant, 1 = Somewhat Relevant, 2 = Largely Relevant, and 3 = Extremely Relevant).

For the first measure, seven SMEs each completed 36 rankings mapping the relevance of all knowledge areas and skills defined in the air-to-air MECs (Colegrove & Alliger, 2002) to our benchmark scenarios.

For the second measure, four SMEs each completed 1,739 ratings of the relevance of all conceptual performance measures to the air-to-air knowledge areas and skills defined in the air-to-air MECs (Colegrove & Alliger, 2002).

The final set of ratings mapped the relevance of objective conceptual performance measures (developed as part of a Performance Measurement Workshop) to objective PETS measures. For this task, seven F-16 SMEs each completed 2,194 ratings.

ANALYSES

We designed the analyses to identify the correspondence between objective performance measure and subjective evaluations provided by SMEs.

Step One: In step one we calculated the average for the ratings for MEC knowledge areas and skill relevance to benchmark scenarios (measure 1) across the SMEs. These ratings provided the basis for organizing those skills and areas of knowledge based on relevance to the benchmark scenarios.

Step Two: In this step, we combined the ratings identifying the degree to which the MEC knowledge and skills are involved in the benchmark scenarios with the ratings evaluating the relationship between the MEC knowledge and skills and the conceptual performance measures. The new scores represent the relationship of the MEC knowledge and skills to the conceptual performance scores, weighted by the degree to which the benchmark scenarios capture each of the MEC knowledge and skill areas. The sum for each PETS conceptual measure is computed to represent the degree to which each conceptual measure is influenced by the MEC knowledge and skills trained on the benchmark scenarios.

Step Three: Based on the SME subjective assessments step three determined the degree to which each metric influences benchmark scenarios. We multiplied the scores derived in step two by the ratings from the mapping between the conceptual measures and the metrics (step one). The resulting values represent the relationship between the conceptual measures and the metrics, weighted by the degree to which those measures would be trained on benchmark scenarios. Finally, these values were summed across the conceptual measures for each metric, resulting in a single value for each metric.

Step Four: Step four identified the PETS performance measures that improved across DMO training research week. We entered the metrics in the three areas of interest into the data set with the value that represented the proportion of improvement on the metric over the week. Improvement is defined as an increase or decrease in the metric, depending on the desired outcome (e.g. "shortest distance of a striker to base" showed improvement by a percent increase in that distance).

Step Five: In step five, we computed Pearson product-moment correlation coefficients between the objective performance measures from training weeks and the scores for MEC knowledge areas and skills involved in benchmark training, according to subjective evaluations.

RESULTS

For the analysis of the ratings relating MEC knowledge areas and skills to the benchmark scenarios (computed in step 1) the average knowledge rating for the benchmark scenarios was 2.45, with a standard deviation of 0.50. The average skill rating for the benchmark scenarios was 2.66, with a standard deviation of 0.30. The SMEs rated both the MEC knowledge area and skills with average ratings between approximately 1.5 and the maximum of 3. This range in scores indicates the high level of relevance of the benchmarks to the knowledge and skills necessary for pilot readiness, while still being able to discriminate between more and less relevant skills and areas of knowledge; Table 1 presents the top five MEC knowledge areas and skills.

Table 1. Top five MEC Knowledge Areas and Skills

Top 5 MEC Knowledge Areas
1. Mission Objectives
2. Threat Capabilities
3. Communication Standards
4. Commit Criteria
5. Formation

Top 5 MEC Skills
1. Builds Picture
2. Listens
3. Multitasks
4. Radar Mechanization
5. Sorts Targets

The second step generated scores that provided an indication of the relevance of each PETS conceptual measure to the benchmark scenarios. We computed an average score for knowledge areas and skills for each conceptual performance measure. There are 12 MEC knowledge areas and 24 MEC skill areas. The average score for MEC knowledge across the conceptual performance measures is 1.89, with a standard deviation of 0.88. The average score for MEC skills across the conceptual performance measures is 2.42, with a standard deviation of 1.07. There are 44 conceptual performance measures in this study. Table 2 illustrates the top five conceptual performance measures influenced by MEC knowledge areas and skill for the benchmark scenarios.

Table 2. Top five Conceptual Performance Measures for MEC Knowledge Areas and Skills

Top 5 Conceptual Performance Measures for MEC knowledge
1. How close red came to point/area/HVAA
2. Number of visual merges with second red within factor range
3. Fly into frag
4. Air-to-air shot measures
5. How many times painted by red air radar

Top 5 Conceptual Performance Measures for MEC skills
1. Quality of communications
2. Mutual support
3. Number visual merges with second red within factor range
4. Percent of red air targeted by targeting range
5. Percent of red air detected by min targeting range

During the third step, we calculated a weighted score representing the degree to which each of the PETS performance measures should improve based on the SME subjective assessments. To identify the degree to which each of the PETS metrics included in the current study would change based on subjective assessments, the relevance of each of the metrics to training benchmark scenarios. The average knowledge score across PETS metrics for this step was 2.09, with a standard deviation of 0.38. The average skill score across PETS metrics for this step was 3.00, with a standard deviation of 0.48.

In the fourth step, we identified seventeen performance measures from PETS to include in the current analyses. We extracted the percent improvement for each metric, based on change over the week to the end of the training week. Table 3 shows the top five and bottom five rank ordered measures.

Table 3. Top five and bottom five metrics showing improvement

Top 5 Metrics
1. Bombers killed before reaching base
2. Average N-Pole Exposure Time
3. Bombers reaching base
4. MAR-1 time for team
5. MAR time for team

Bottom 5 Metrics
5. MOR time for team
4. Slant range to target (AAMRAM) at launch
3. 2D range to target (AAMRAM) at launch
2. Proportion of all threats killed
1. Proportion of Viper shots resulting in kill

The final step compared the degree to which pilots improved on different objective performance measures with the anticipated improvement on the measures, based on the subjective SME assessments. A correlation between the scores from MEC knowledge areas and the percent improvement was not significant, $r(15) = 0.23$, n.s. The correlation between the scores from MEC skills and the percent improvement was not significant, $r(15) = 0.20$, n.s. In order for a correlation to be significant with 15 degrees of freedom the value of the coefficient would need to be .48.

DISCUSSION

Our findings provide preliminary support for further development of the process presented here. Identifying the areas in which subjective and objective performance measurements are most effective and efficient offers a powerful tool for developing and refining training programs. Additionally, the correspondence between subjective and objective performance measures that we report here would enable instructors to select and integrate objective performance measures into training. For example, if an instructor sees that a pilot is not improving on certain objective performance metric, they can use the correspondence to know which MEC skills and knowledge should areas should be remediated in training. Additional investigations will refine the process to provide a more rigorous closed-loop, adaptive training process.

The lack of significant correlations between the subjective scores and the objective improvements should not be interpreted as a lack of evidence for the process. Although the correlations were not found to be significant, only 17 PETS metrics were used in the current study, providing few degrees of freedom. The correlation coefficients, though in the range of small relationships, were both in the correct direction and represent small effect sizes.

In addition to the small number of metrics included in this study, this is the first time that this rating system for mapping measurement frameworks has been used in this environment and is still in the testing phase of the development process. The knowledge, skill, and performance mappings were done with a small sample size to provide enough data to validate the process. An

increase in the number of SMEs providing ratings for mappings may provide for sensitive measures, decreasing the variability and improving the relationship between the objective and subjective performance measures.

Although the findings could have been stronger for validating the relationship between objective and subjective performance measures, the results of the process do provide a strong framework for building an understanding of the relationships. The use of objective performance data in the training environment will ultimately be limited on the ability of instructors and trainees to disseminate and understand the feedback from the objective measurement systems.

The process presented in the current framework can be used to develop more sophisticated competency-based training environments. Furthermore, once the process explored in this study is validated the metric can be used as an assessment tool in an adaptive training environment. Future research might investigate the full range of available objective performance metrics and the impact of system fidelity on the mapping process. Finally, the next goal of the current research will be to integrate this work as an additional tool for enhancing training environments.

ACKNOWLEDGEMENTS

REFERENCES

Colegrove, C. M., & Alliger, G. M. (2002, April). Mission Essential Competencies: Defeining Compat Mission Readiness in a Novel Way. *Paper presented at NATO RTO Studies, Analysis and Simulation (SAS) Panel Symposium.* Brussels, Belgium.

Crane, P. M., Robbins, R., & Bennett, W. J. (2000). Using Distributed Mission Training to Augment Flight Lead Upgrade Training. *2000 Interservice/Industry Training, Simulation and Education Cnference (I/ITSEC) Proceedings.* Orlando, FL: National Security Industrial Association (AFRL-HE-AZ-TR-2000-0111, ADA394919). Proj 2743. F41624-97-D-5000. Mesa, AZ: L3 Communications.

Cushman, W. H., & Rosenbery, D. J. (1991). *Human factors in product design.* Amsterdam: Elsevier.

Denning, T., Bennett, W. Jr., & Crane, P. M. (2002). Mission Complexity Scoring in Distributed Mission Training. In 2002 *Proceedings of the Interservice/Industry Training, Simulation and Education Conference (I/ITSEC).* Orlando, FL: National Security Industrial Association.

Krusmark, M., Schreiber, B. T., & Bennett, W. J. (2004). *The Effectiveness of a Traditional Gradesheet for Measuring Air Combat Team Perfomrance in Simulated Distribute Mission Operations.* (AFRL-HE-AZ-TR-2004-090). Air Force Research laboratory, AZ: Warfighter Readiness Research Division.

MacMillan, J., Entin, E. B., Morley, R., & Bennett, W. Jr. *Measuring team performance in complex and dynamic military environments: The SPOTLITE method.* Manuscript in preparation.

Rowe, L. J., Gehr, S. E., Cooke, N. J., & Bennett, W. J. (2007). Assessing Distributed Mission Operations Using the Air Superiority Knowledge Assessment System. *Hman Factors and Ergonomics Annual Meeting.* Baltimore, MD.

Rowe, L. J., Schvaneveldt, R. W., & Bennett, W. J. (2007). Measuring Pilot Knowledge in Training: The Pathfinder Network Scaling Technique. *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC).* Orlando, FL: National Security Industrial Association.

Schreiber, B. T., & Bennett, W. J. (2006). Distributed Mission Operations Within-Simulator Effectiveness Baseline Study: Summary Report. *(AFRL-HE-AZ-TR-2006-0015-Vol I). Air Force Research Laboratory, AZ: Warfighter Readiness Research Division* .

Schreiber, B. T., & Bennett, W. J. (2006). *Distributed Mission Operations Within-SImulator Training Effectiveness Baseline Study.* Mesa AZ: Air Force Research Laboratory, Warfighter Training Resaerch Division.

Schreiber, B. T., Gehr, S. E., & Bennett, W. J. (2006). Distributed Mission Operations Within-Simulator Training Effectiveness Baseline Study: Real-Time and Blind Expert Subjective Assessments of Learning. *AFRL-HE-AZ-TR-2006-0015-Vol II* . Air Force Research Laboratory, AZ: Warfighter Readiness Research Division.

This page intentionally left blank.

Student Flight Instructor Competencies

Patricia C. Fitzgerald, Dee H. Andrews
Air Force Research Laboratory
Mesa, AZ
Patricia.Fitzgerald@mesa.afmc.af.mil,
Dee.Andrews@mesa.afmc.af.mil

Brent Crow
Consortium Research Fellows Program
Mesa, AZ
Brent.Crow@asu.edu
Merrill R. Karp, Jim Anderson
Arizona State University
Ron.Karp@asu.edu, J.Anderson@asu.edu

ABSTRACT

The research literature addresses a variety of questions concerning flight instructor training, however, more research is needed to elucidate the instructional competencies associated with successful instruction in this critical field. This paper presents observational research to identify flight instructor competencies and patterns of instructional behavior. Flight instructor behaviors were defined in a computer-based observational tool that allows behaviors to be logged. Seventeen Certified Flight Instructor Instrument (CFII) students were videotaped as they were instructing Instrument flight students on a flight simulator. The researchers coded the student's behaviors using an observational data collection tool. Observed behavioral patterns are presented. The identification of critical instructional competencies during training and the use of the computer-based behavior logging tool in training flight instructors is discussed. Follow-on studies to further investigate methods of enhancing instructor performance are presented.

ABOUT THE AUTHORS

Patricia C. Fitzgerald is a Research Psychologist with the Warfighter Readiness Research Division in the Human Effectiveness Directorate of the Air Force Research Laboratory. In addition to conducting a variety of training research studies, she is the Program Manager for the Improved Performance Research Integration Tool (IMPRINT) Training Algorithm Enhancement program. She served as an Information Systems Specialist for 20 years before beginning her career with the Air Force. Ms. Fitzgerald holds a Master of Science degree in Aviation Human Factors and a Bachelor of Arts degree in Psychology.

Dr. Dee H. Andrews is the Directorate Senior Scientist (ST) for the Human Effectiveness Directorate of the Air Force Research Laboratory (AFRL). As a Senior Scientist (ST), Dr. Andrews is the AFRL's principal scientific authority for training research. Dr. Andrews' responsibilities include sustaining technological superiority for training by planning and conducting theoretical and experimental studies. He also is responsible for mentoring and developing dedicated technical staff to assure quality in training research, and he represents AFRL in training research matters to the external scientific and technical community.

Brent Crow is a Research Fellow with The Consortium Research Fellows Program. He is working on his Master's degree in Aviation Human Factors at Arizona State University and expects to complete his degree requirements in December 2008. Mr. Crow is a Certified Flight Instructor and received his Bachelor of Science degree in Professional Flight at Arizona State University in 2006.

Merrill R. (Ron) Karp is a Professor of Practice at Arizona State University and has been a member of the Aeronautical Management Technology Department faculty for 14 years. Dr. Karp is a retired colonel in the U.S. Air Force and was the Commander of the F-4G Wild Weasel Wing during the first Gulf War. Dr. Karp earned his Ph.D. at Walden University in 1996, with a specialization in Aviation Education and Training; he received his Masters degree from Central Michigan University and his undergraduate BS degree from the Aeronautical Management Technology Department at Arizona State University in 1967.

Jim Anderson serves as a lecturer and the CRJ200 Flight Training Device (FTD) Coordinator in the Aeronautical Management Technology Department in the College of Science and Technology at Arizona State University. Prior to retiring after 16½ years of flying with Southwest Airlines, Jim had a 20 year career in the United States Air Force and flew a variety of aircraft. Jim has logged over 19,000 hours throughout his flying career. He retired from the Air Force as a Lieutenant Colonel and has a BS in Military Arts and Sciences from the USAF Academy and a MA in Business Management from Central Michigan University.

Student Flight Instructor Competencies

Patricia C. Fitzgerald, Dee H. Andrews
Air Force Research Laboratory
Mesa, AZ
Patricia.Fitzgerald@mesa.afmc.af.mil,
Dee.Andrews@mesa.afmc.af.mil

Brent Crow
Consortium Research Fellows Program
Mesa, AZ
Brent.Crow@asu.edu
Merrill R. Karp, Jim Anderson
Arizona State University
Ron.Karp@asu.edu, J.Anderson@asu.edu

INTRODUCTION

Pilot instructing was first done by the Wright Brothers as they taught themselves to fly, and then taught their early customers. From that time forward hundreds of thousands of aviators have served as flight instructors (civilian term) and instructor pilots (military term). Not surprisingly, after the early days of flight instruction the instructional role has always fallen to aviators who have a good bit of aviation experience. Instructors are usually chosen because they have shown their skill at aviation. However, as is the case with university teaching, skill at instructing is not necessarily a major criterion for being selected. It is not typically known who will be a good flight instructor until a candidate has tried to instruct. The literature review below will show that after all these years the aviation community has little in the way of analytical evidence that informs those responsible for instruction about how best to select or train flight instructors. It is fair to say that flight instruction is still far more art than it is science.

Our literature search has revealed few studies that examine analytically or empirically the question, "What makes a good flight instructor?" In addition, we have found few research based articles that ask, "How can flight instructors be better prepared?" While the military has a number of quality courses for preparing instructors, their curricula do not have a substantial theoretical or analytical base. Pedagogical skills are taught, but providing instructor candidates and their instructors with a well researched set of models for quality instruction is not possible because such research is not available.

We undertook this research with the goal of developing instructor guidelines based on sound instructional theory and analytical data. We desired to provide a set of valid guidelines that could be used by new instructors with behaviors that would result in better teaching. We desired that these modeled behaviors could be used in simulators and aircraft cockpits. Rather than base these instructional behavioral models only on subject matter expert opinion we felt it important to model excellent instructor behaviors so that new instructors could attempt to emulate the excellent instructors' approach to teaching.

LITERATURE REVIEW

Current Civilian Instructor Pilot Training

The Federal Aviation Regulations 14 CFR Part 61.181 outlines the eligibility, aeronautical knowledge, and flight proficiency requirements for flight instructor applicants (FAA, 2005). Prior to becoming a flight instructor, applicants must pass two multiple choice written exams: one on the fundamentals of instructing and another on general flight knowledge. Recent research suggests that most applicants memorize the correct answers (Casner, Jones, & Irani, 2004). Nevertheless, flight instructor applicants are verbally quizzed by a Designated Pilot Examiner during the oral exam which they must pass as well. According to the Practical Test Standards, the Designated Pilot Examiner has the responsibility for determining that the applicant meets acceptable standards of teaching ability, knowledge, and skill required in each of the tasks found in the Practical Test Standards (FAA, 2002). Most of the tasks in the Practical Test Standards require that the applicant demonstrate instructional knowledge by being capable of using the appropriate reference to provide the application or correlative level of knowledge of a subject, procedure, or maneuver. The applicant must also follow the recommended teaching procedures and techniques explained in the Aviation Instructors Handbook (FAA, 2002). This means that the instructor applicant comes prepared with a lesson plan outlining the objectives, elements, and completion standards for the lesson they are going to teach their Designated Pilot Examiner. Generally, a

flight instructor will help their instructor applicant or student develop a lesson plan, and practice giving the lesson to their instructor. Unfortunately, this may be the only instance in which the applicant may use a lesson plan, as many flight instructors do not create lesson plans prior to scheduled flights or ground training. Finally, the applicant must satisfactorily pass a practical test on the areas of operation listed in 61.187(b) and must once again demonstrate instructional knowledge in the elements and common errors of a maneuver or procedure (FAA, 2005). A typical flight training session for an instructor applicant in order to prepare for the above practical test requires that the student instructor practice instructing on their instructor, who will play the role of both mentor and student.

Shortfalls of the Current Flight Instructor Certification Process

The method described above for determining flight instructor competency is insufficient. As Machado (2005) described, "It is better to spend three years looking for a good instructor, than spend three minutes with a bad one". Although the FAA has a stringent certification process, ineffective instructors occasionally progress to student instruction (Wright, 2003). Further research will be necessary to mitigate this problem. Perhaps the reason is because flight instructor applicants can easily pass two written tests, teach a few lessons to their flight instructor, and show their teaching ability to a Designated Pilot Examiner who has a widely varying view of competency (Hunt, 2001). In this example, a flight instructor applicant has only been teaching to an audience that already knows the relevant information to a level higher than the applicant. Instructors know what examiners are looking for, and therefore, often teach their student to just pass the test, robbing them of the skills, knowledge, and attitudes necessary for daily flight (Hunt, 1997; Lintern, 1995; Moore, Lehrer, & Telfer 1997). The maneuvers required on the practical test do not have content or criterion validity (Blickensderfer, Schumacher, & Summers 2007). Role-playing as an instructor toward their designated examiner during the practical exam and to their instructor during training is confusing and unrealistic. This is evident in research done by Henley (1995) in Canada and Australia, and in the United States, it is understood by the FAA to be taking place (Wright, 2003).

Further research in the field of aviation instruction competencies would yield a better understanding of the requirements for training instructors. It may be valuable to consider the research of the Committee on techniques for the Enhancement of Human Performance which discovered that performance during training is an unreliable predictor of learning real world tasks (Druckman & Bjork 1994). Instructor applicants are sure to find that teaching their flight instructors and Designated Examiners is a simple task since they already understand the material. However, when given the task of training a new student, questions remain concerning actual instructional effectiveness.

Flight Instructor Training Research

Although the training of pilots has received a great deal of empirical research attention over the years, a review of the literature revealed little in terms of addressing the multiple factors associated with good flight instruction in military or civil aviation. A number of researchers, however, have addressed specific issues associated with flight instruction.

One line of investigation addresses pilot performance rating by instructors. In one study, Mulqueen, Baker, and Dismukes (2002) investigated the rating behaviors of commercial flight instructor's evaluations of pilots' technical and Crew Resource Management (CRM) skills in a flight simulator scenario. The goal of this effort was to assess the extent to which instructor ratings of pilot performance were accurate and reliable. Results indicated that participants had more difficulty assessing CRM skills than technical skills and that rating inconsistencies existed, suggesting the need for rater training programs to address these issues. In another study, Greenwood, Holt, and Boehm-Davis (2002) evaluated the efficacy of two training interventions to enhance inter-rater reliability among airline instructor pilots. One focused on conceptual knowledge while the other focused on procedural knowledge. The findings indicated that while participants in both training tracks experienced increased learning of concepts and procedures, participants in the procedural track reported higher levels of pre- and post-workshop knowledge. The authors conclude that the use of multiple index profile inter-rater reliability led to improved reliability of groups of raters and also that evaluators/instructors that lack a statistical background could indeed use a procedurally-based evaluation system.

In a study of the use of facilitation by instructors in debriefing following Line-oriented flight training simulator sessions, the techniques utilized by the flight instructor were investigated (Dismukes, Jobe, &

McDonnell, 1997). In this study, the ways in which commercial flight instructors facilitated crew self-reflection and self-assessment following a simulator flight were explored. While a focus on crew performance was evident, instructors were more likely to emphasize the positive events of the session rather than the aspects that needed improvement. Furthermore, the sessions were marked by frequent instructor questions to stimulate discussion. Included in the behaviors evident among the instructors who facilitated the debriefings effectively were: the use of questions that promoted self-analysis, appropriate silence, active listening, and follow-up questions. Interestingly, when effectiveness of facilitation skills was analyzed, a bi-modal distribution emerged, with a large group of instructors in the "good" to "very good" range and another large group in the "marginal" range. These results strongly suggested the need for facilitation training that includes hands-on practice and mentoring from instructors experienced in facilitation techniques. In another study, Beaubein and Baker (2003) found that there were no differences between team and instructor-led flight debriefings. Although the researchers reported that these debriefing methods were equally effective, further research was recommended to investigate ways to improve debriefing effectiveness.

A number of studies concerning flight instructor education were conducted by Irene Henley and her colleagues. In one study, a survey was conducted to elucidate the factors associated with the development and evaluation of flight instructors (Henley, 1991). Results of this survey showed that flight instructor training is highly influenced by traditional methods of flight instruction such as rote memorization and modeling other instructors. Deficiencies noted were a lack of identifiable instructor competencies and insufficient training in instructional methods. In another survey-based study, Henley (2001) discovered that the main hindrance to student learning in aviation education was their instructor, the very person who should be focused on promoting student learning. Specifically, flight instructors caused the most stress for flight students and were called, "the weakest link" in flight training (Henley, 2001).

These investigations provide valuable insight into some of the key factors associated with effective flight instruction. Gaining a greater understanding of the *behavior patterns* that are related to effective instruction during flight, however, is the goal of this research program.

Instructor Competencies

In an effort to ensure that personnel have the requisite skills to perform their jobs, employers are increasingly relying on the use of professional competencies in selection and hiring decisions, performance assessment, and training programs. The Department of Education for example, sponsored a program to develop an Instructor Competencies Assessment Instrument based on previously identified adult educator competencies (Sherman, Dobbins, Crocker, & Tibbett, 2002). This instrument is used in a variety of adult educational settings.

The International Board of Standards for Training, Performance and Instruction (IBSTPI), in cooperation with the Association for Educational Communications and Technology, conducted an empirical study to determine the competencies associated with effective instruction (Klein, Spector, Grabowski, & de la Teja, 2004). The use of the IBSTPI competencies for the current study will be discussed further in the methods section of this paper.

Observational Data Collection

Observing participants and collecting data in a natural setting often pose a number of challenges. It is widely accepted by research practitioners that the mere act of observing behavior may in fact change that behavior. While it is difficult to determine the extent to which this occurs in any setting, researchers try to minimize their impact on behavior in a number of ways. Using a recording device is one way to minimize the effects of the observer.

How to collect the data may pose additional challenges. It may be difficult to interpret, process, and record behavioral data during fast-paced human interactions. If the behaviors of interest are few, it may be possible to effectively collect the data in real time. The complexity of the environment, along with the number of observed participants, however, quickly exposes the limits of the researcher.

In an early attempt to automate observational data collection, a typewriter was modified to record the interactions of teachers and students in a classroom setting. (Young & Wadham, 1975). The Time Interval and Categorical Observation Recorder (TICOR) was designed to facilitate the coding of behavioral data and allowed the capture of the duration of the behavior. This system allowed researchers to ascertain patterns of behavior between the student and the instructor, leading to the ability to

conduct cause-and-effect analyses. The system was devised so that the recorder could enter a behavior, along with the quality of the behavior with as little as three keystrokes. For example, an incorrect learner response would require the researcher to enter R-. Because time and duration data were collected, researchers could then analyze patterns in the behaviors of the students and the teachers. Although this was a very innovative at the time, a number of more sophisticated computer systems have been developed to collect behavioral data. One such system was selected for this Instructor Pilot Training study and will be discussed in greater detail below.

METHODS

Development of Behavioral Assessment Tool

The research effort discussed in this paper is the most recent in a series of studies investigating instructor pilot behaviors, leading to the development of a tool to aid in training. Working with instructors at Arizona State University's aviation department, The Air Force Research Laboratory identified instructor pilot behaviors that facilitate student learning. Initial instructor behaviors were derived from the instructor competencies research conducted by IBSTPI (Klein, Spector, Grabowski, & de la Teja, 2004). A comprehensive set of behaviors was identified in this research effort, and from that set, a subset that was most relevant in the aviation setting was derived. The reason for limiting the number of behaviors for the current effort was twofold. First, not all of the behaviors identified by IBSTPI are used in one-on-one instruction. For example, improving professional knowledge and skills is undoubtedly imperative for instructors in any field; however, the behaviors associated with this competency would be difficult to quantify in the context of the present study. Secondly, the investigators felt that it was more important to focus on the most relevant behaviors for one-on-one instruction in typical aviation instructional experiences. Specifically, a great deal of instructor-student interaction takes place in a simulator, aircraft, or a briefing/debriefing setting. Focusing on the key behaviors in these settings would result in a more useful tool for instructors to use in simulator and cockpit training.

To further refine our list of behaviors, experts in the field then supplemented the initial behavior set to include several aviation-specific behaviors. For instance, if done appropriately, assisting a student when workload limits are exceeded facilitates learning. Depending on the student's level of

proficiency, events for which a student does not have experience may interfere in the student's ability to absorb the objectives of the training session. Instructor intervention in events that are not relevant to the session allows the student to focus on flight objectives. Conversely, if an instructor intervenes too often, the student may become over-reliant on the instructor, and may not learn the important points of the lesson. Capturing such behaviors was imperative for accurate assessment of flight instructor teaching behavior.

The behavior set was then entered into a data collection software package. A behavioral analysis research tool, Noldus Observer XT facilitates coding of the behaviors of one or more participants in an observational research setting (The Observer XT, n.d.). Once the behaviors are entered into the system, the patterns of behavior may be represented on a chart (figure 1). These charts may be used by instructor pilot trainees to gain a better understanding of the behaviors they used in a training session. Furthermore, if learner behaviors are also coded, the ways in which students respond to instructor actions may also be assessed. Over the course of several semesters, data were collected during training sessions on a simulator. The researchers and flight instruction experts assessed and refined the behaviors under investigation. The results and findings of these previous efforts led to the development of the methods for the present study.
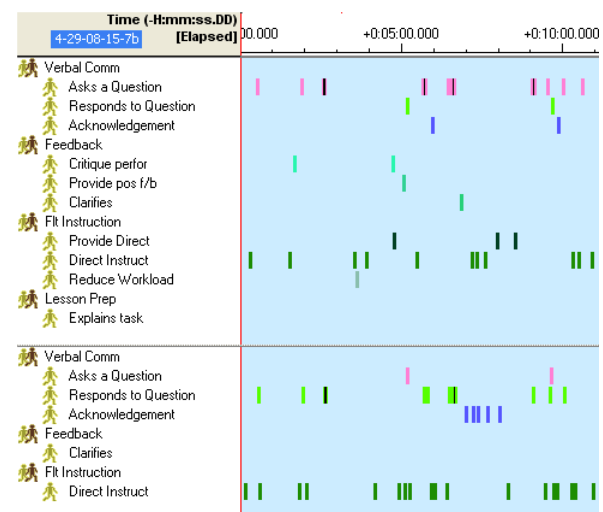

Figure 1. Noldus observed behavior chart.

The Present Study

During the spring 2008 semester at ASU, 17 flight instructor trainees were recorded while instructing

instrument flight students on a flight training device. These instructors-in-training hold a commercial certificate with an instrument rating and are working toward obtaining their Certified Flight Instructor (CFI) certificate. The instrument students are working on obtaining, or currently have, a private pilot certificate, and are beginning their ground training in instrument flight.

The equipment used for the training sessions consisted of an ELITE PI-126 Personal Computer Aviation Training Device (figure 2).

Figure 2. Personal Computer Aviation Training Device.

Using this device as a training platform, the student instructors taught instrument training skills such as holding, tracking a Non-Directional Beacon (NDB) or Very high frequency Omni-Directional Range radio (VOR), or a segment of an instrument approach. Scenarios were also flown in which the student instructor and instrument student had to fly an instrument approach with air traffic control. The researchers observed 19 sessions. The video recordings were then coded by the researchers using the observational software discussed above. For the current study, the 22 behaviors previously defined were used with each behavior given a keystroke assignment (see figure 3).

The behavior "Ask a Question" for example, was given the keystroke "aq," so that when watching the video recording, each behavior observed could be coded in real-time by a simple keystroke. After each observation, the observational tool provided the number of times each behavior was coded in the observation, as well as other descriptive information. Since each observation was 15-40 minutes in length,

the researchers used rate per minute (RPM) data for each of the behaviors so that time was not a confounding factor in our analysis. Not every behavior was analyzed, as some did not occur, or occurred too rarely, to prove meaningful. Any behaviors that occurred fewer than 5 times were excluded from the analyses. Thus, 9 behaviors proved useful for the study. Behavioral data were then used to generate observed behavior charts, depictions of the occurrence of all behaviors over time.

Figure 3. Flight Instructor Behaviors.

RESULTS

The observed behavior charts displayed a great deal of variation among the student instructors. Refer to table 1 for a chart depicting the rate per minute (RPM) of each of the behaviors. Although conclusions may not be drawn because we do not have performance data, it is interesting to note the large differences in instructor behaviors across the different observations. Some instructors talk to their students nearly continuously while others seldom talk at all. In the sessions observed for this study, the more behaviors the student instructor exhibited, the more behaviors the student exhibited ($r = .685$, $p < .01$). The three most frequently occurring behaviors

were: direct instruct (e.g., providing a truism, such as "we are at 2000 feet"), provide direct (e.g., provide a command, such as "descend to 2000 feet"), and ask a question. The three least common of our selected behaviors were: clarifies, reduce workload, and explains task.

Table 1. Student Instructor Behavior Rates

	Rate Per Minute (RPM) Across All Student Instructor Observations				
	Average	Minimum	Maximum	Std. Dev.	Sum
Acknowledgment	0.33	0.00	0.62	0.16	3.28
Responds to Question	0.38	0.00	1.02	0.29	7.18
Asks a Question	0.53	0.00	1.38	0.44	10.14
Explains Task	0.20	0.00	0.57	0.18	3.74
Reduce Workload	0.16	0.00	0.42	0.13	2.98
Direct Instruct	1.08	0.34	2.37	0.52	20.47
Provide Direct	1.03	0.11	2.16	0.55	19.51
Clarifies	0.11	0.00	0.39	0.12	2.04
Provide Postive F/B	0.33	0.00	1.21	0.30	6.33

DISCUSSION

It is anticipated that the tool being developed for this research program will provide a valuable resource during the training of future flight instructors in civil and military aviation. Although video tapes for reviews of instructional behavior are seldom used during debriefings, one could argue that doing so could enhance self- and instructor-assessment. The inclusion of the tool being developed through this research program will provide valuable information on the frequency and distribution of instructional behavior. Furthermore, this tool will enable student instructors to evaluate the ways in which their students respond to instruction.

Figures 4 and 5 depict the behavior patterns of two of the instructors that participated in the study. Coded instructor behaviors appear above the line; student behaviors are represented below the line. The behavioral patterns depicted in figure 4 suggest that the instructor is proactive, periodically asking the student questions in order to determine their level of understanding. The student responds to questions and asks some of their own. The observed behavior chart also reveals that this instructor offers positive feedback to the student and clarifies information at various points in the simulator session. It is also useful to note the behaviors that did not appear in the observed behavior chart. For instance, critiques were not provided, and the instructor did not intervene or

reduce workload during the simulator session. Depending on the circumstances of the flight, the presence or absence of these behaviors may be meaningful, potentially prompting discussions concerning instructional improvements.

In contrast, the instructor's behavior pattern depicted in figure 5 shows that this instructor exhibited much less activity. This instructor was passive, asking no questions and only responding to a few posed by the student. This is not to say that one of these instructors is better than the other; rather, these differences can be easily viewed by a student instructor who can make the determination based on the situation, depending on what was more appropriate for the session.

Future Research

The researchers have many suggestions for future research. During the next semester, research plans include obtaining model behavior patterns from expert flight instructors. These behavior patterns are expected to be useful guides for student instructors in developing their instructional techniques. These patterns are not intended to be a prescription for effective instruction; rather they offer alternatives for different approaches to instruction.

Since flight training is not one-size-fits-all, instructors must be able to tailor their instruction to meet the educational needs of the student. This research into instructional behavior patterns may shed light on the effectiveness of different techniques. Commonly, beginning student pilots need a great deal of interaction with their instructors, whereas checkride-ready students require significantly less. By assessing flight specific behavior patterns of both student and instructor, adjustments could then be made to achieve the optimal flight training environment.

Finally, recording student instructors on a simulator has been useful for developing our methods, but we intend to take this idea into the cockpit to observe certificated flight instructors teaching actual students to become flight instructors.
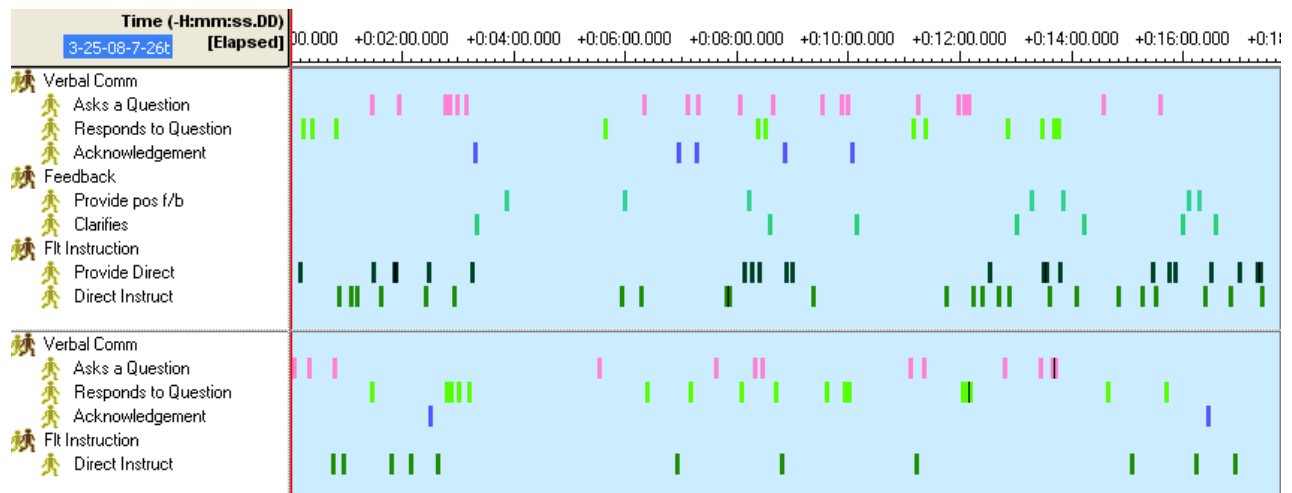
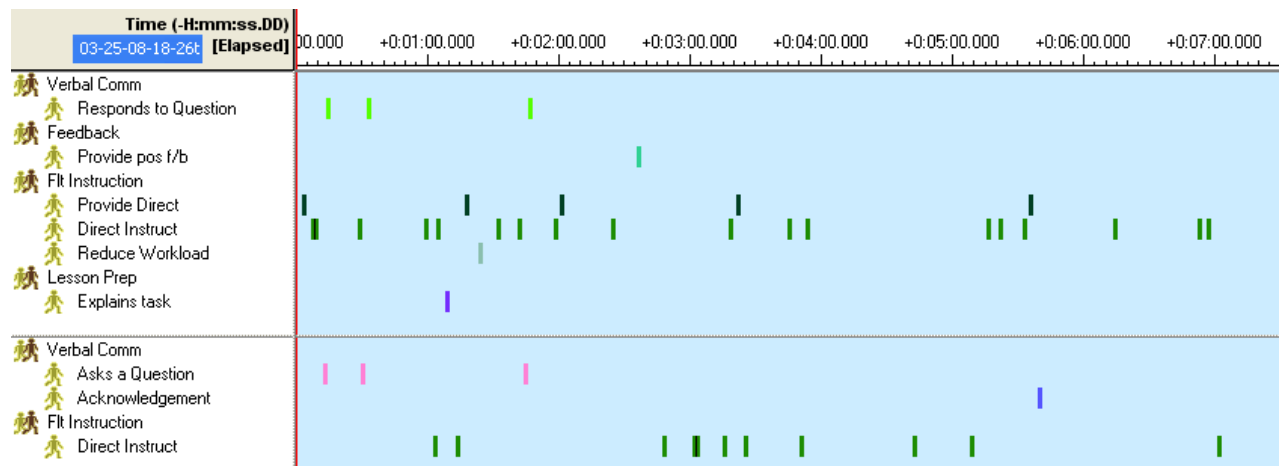**Figure 4. Instructor 1 behavioral pattern.**

Figure 5. Instructor 2 behavioral pattern.

ACKNOWLEDGEMENTS

REFERENCES

Beuabein, J. M. & Baker, D. P. (2003). Post-training feedback: The relative effectiveness of team-versus instructor-led debriefs. In *Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 2033-2036). Denver, CO.

Blickensderfer, E. L., Schumacher, P., & Summers, M. (2007). Learner centered debriefing in general aviation training: Questions from the field and answers from research. In R. Jensen (Ed.), *Proceedings of the 14th International Symposium on Aviation Psychology* (pp. 45-50). Dayton, OH.

Casner, S. M., & Jones, K. M., & Irani, H. (2004). *FAA pilot knowledge tests: Learning or rote memorization?* (Technical Report No. NASA/TM-

2004-212814). Moffett Field, California: NASA Ames Research Center.

Dismukes, R. K., Jobe, K. K., & McDonnell, L. K. (1997). LOFT debriefings: An analysis of instructor techniques and crew participation (NASA Technical Memorandum 110442 DOT/FAA/AR-96/16).

Druckman, D. & Bjork, R. A. (1994). In D. Druckman & R. A. Bjork (Eds.), *Learning, remembering, believing: Enhancing human performance*. United States of America: National Academy Press. Retrieved October 30, 2006 from http://books.nap.edu/openbook.php?record_id=2303&page=R1

Federal Aviation Administration (FAA). (2002). In FAA (Ed.), *Flight instructor practical test standards* (FAA-S-8081-6BS ed.). Oklahoma City, OK: FAA.

Federal Aviation Regulations (FAR), 40CFR U.S.C. 61.181 (2005).

Greenwood, D. M., Holt, R. W., & Boehm-Davis, D. A. (2002). Training airline instructor pilots to evaluate air crew performance: Inter-rater reliability training. *International Journal of Applied Aviation Studies*, 2(2), 37-56.

Henley, I. (1991). The development and evaluation of flight instructors: A descriptive survey. *The International Journal of Aviation Psychology*, 1(4), 319-333.

Henley, I. (1995). Flight instructor test – Is it a valid and reliable arssessment of an irnstructor's teaching competence? In R. Jensen (Ed.), *Proceedings of the 8th International Symposium on Aviation Psychology*. Dayton, OH.

Hunt, G. J. F. (2001). On becoming professional: Research issues for professional aviation accomplishment. In R. Jensen (Ed.), *Proceedings of the 11th International Symposium on Aviation Psychology*. Dayton, OH.

Hunt, L. M. (1997). The specification of competencies for the professional aviation instructor. *Proceedings of the 9th International Symposium on Aviation Psychology* (pp. 1223-1227). Dayton, OH.

Klein, J.D., Spector J.M., Grabowski, B., and de la Teja, I. (2004). *Instructor Competencies: Standards for Face-to-Face, Online, and Blended Settings*. Greenwich, CT: Information Age Publishing.

Lintern, G. (1995). Flight instruction: The challenge from situated cognition. *International Journal of Aviation Psychology*, 5(4), 327-350.

Machado, R. (2005). *Flight training interview*. Retrieved October 20, 2006 from http://www.rodmachado.com/Articles/Interview-1.htm

Moore, P. J., Lehrer, H. R., & Telfer, R. A. (1997). Instructor perspectives on learning in aviation. *Proceedings of the 9th International Symposium on Aviation Psychology* (pp. 1252-1255). Dayton, OH.

Mulqueen, C., Baker, D. P., & Dismukes, R. K. (2002). Pilot instructor rater training: The utility of the multifacet item response theory. *The International Journal of Aviation Psychology*, 12(3), 287-303.

Sherman, R., Dobbins, D., Crocker, J., & Tibbetts, J. (2002). Instructor competencies assessment instrument. Retrieved June 4, 2008, from http://www.pro-net2000.org/CM/content_files/70.pdf

The Observer XT. (n.d.). Retrieved June 11, 2008 from http://www.noldus.com/site/doc200401012

Wright, R. A. (2003). Changes in general aviation flight operations and their impact on system safety and flight training. In *Proceedings of the International Air and Space Symposium and Exposition; the next 100 years* (pp. 957-964).

Young, J.R. & Wadham, R. A. (1975). A simplified evaluation program to distinguish different levels of teacher competence and teacher-pupil interaction patterns based on pupil responses. Provo, UT: Brigham Young University. (ERIC Document Reproduction Service No. ED132131).

Observers measure warfighter performance during October 08 training research exercise. Photo by Bruce Liddil.

AOC Training Research Exercise (T-REX) Hits New Heights

The Warfighter Readiness Research Division (711 HPW/RHA) hosted a select group of highly experienced joint warfighters in an October research project. RHA's Air and Space Operations Center (AOC) Training Research Exercise (T-REX) 09-1 investigated immersive training, continuous learning, information simulation, and leading-edge tactics used by the Dynamic Effects Cell (DEC) of the Falconer Combined Air and Space Operations Center (CAOC). Training this team in Distributed Mission Operations (DMO) is a challenge when competing objectives or incomplete scenarios limit the extent participants can exercise the knowledge and skills that are required to be fully mission-ready. In this exercise, the RHA AOC Training Research Team presented an optimized scenario with selected DMO capabilities to focus intensive training on the DEC team. To ensure the highest value of training and knowledge transfer, Mesa's AOC Training Research Team employed DEC subject matter experts from the USAF Warfare Center, Special Warfare Center, and Naval Strike Air Warfare Center.

T-REX 09-1 research objectives targeted improving mission readiness through a continuous scenario containing complex targeting problems exercising the full spectrum of challenges and decisions in both conventional targeting and asymmetric warfare. The team of trainees faced a cell-structured adversary integrated with a local population and an adjacent country's special operations forces. The adversary was technically proficient, expert in counterinsurgency, aggressive, and not constrained by laws of armed conflict. The scenario challenged the team to react quickly and correctly to target adversary warfighting capabilities and support structure while abiding by stringent strategic guidance and coalition country rules.

The combined team led the force in the first trial and analysis of emerging joint command and control doctrine and Improvised Explosive Device network defeat Tactics, Techniques, and Procedures (TTPs) as well as emerging Internet Relay Chat (IRC) employment TTPs. Seven scenario controllers making up the "white cell" created a realistic, information rich environment that set the stage for the eleven-member AOC Dynamic Effects Cell to take on the challenge with a much broader set of tools than conventional dynamic targeting training.

The exercise's detailed scenario and range of available assets provided a forum for training research across the spectrum of solutions, as well as testing integrated kinetic and non-kinetic complementary operations simultaneously. The research targeted effective analysis and debrief of team performance. Subject areas included command and control, systems integration, emerging assessment and debrief tools, communication, white force integration and continuous learning. Analysis by subject matter experts will investigate adherence to draft TTPs and effects on mission performance by examining message effectiveness, chat room use, effects of chat format/content on situational awareness level, and chat information transfer to the Joint Automated Deep Operations Coordination System (JADOCS) collaborative tool. This data reduction and analysis will guide collaborative development and update of emerging and existing after-action reporting tools under development with RHA.

An example of this collaboration is data collected and analyzed on chat room employment using the Chat Information Tracking System (CIFTS). CIFTS was designed and developed under a Small Business Technology Transfer (STTR) effort led by the Air Force Office of Scientific Research in conjunction with RHA

AOC T-REX 09-1 Team. Photo by Bruce Liddil.

and is using techniques in Social Network Analysis (SNA) to measure send and receive patterns. CIFTS also uses SNA visualization tools to give researchers new insights into individual and team performance. T-REX 09-1 marks the first CIFTS trial in exercise conditions.

Another collaborative effort is chat room presence and participation monitoring using a new version of an existing after-action review tool known as CAOC Performance Assessment System (CPAS).

Future Steps for DMO

Researchers develop new training technologies to enhance preparation for Air Force ISR personnel. Photo by Bruce Liddil.

Today's Air Force intelligence personnel work in many different mission areas, with a variety of platforms, and support a broad range of customers often working as geographically distributed teams and with geographically separated customers. Air Force personnel assigned to the Intelligence Surveillance and Reconnaissance (ISR) mission areas can benefit from distributed training constructs like Distributed Mission Operations (DMO) to improve individual and team performance. 711 HPW/RHA is conducting research to enhance the experience and mission readiness of Air Force intelligence personnel through competency-based high-fidelity training methodologies and technologies. 711 HPW/RHA teamed with the Joint System Integration Laboratory (JSIL) to develop a Realistic Training Environment (RTE) proof-of-concept for the Air Force-Distributed Common Ground System (AF-DCGS) Formal Training Unit (FTU) located at Goodfellow Air Force Base. The RTE proof-of-concept system employs the 711 HPW/RHA eXpert Common Immersive Theater Environment (XCITE) to create a synthetic area of operations and utilizes the JSIL developed Air Force Synthetic Environment for Reconnaissance and Surveillance (AFSERS) to simulate ISR platforms. XCITE models adversary, friendly, and neutral computer generated forces. Sensor platforms including the U-2, Predator, Global Hawk, and JSTARS are modeled by the AFSERS simulation. AFSERS provides near-real-time telemetry, fixed frame imagery, video and Moving Target Indicator (MTI) data.

Innovative host processes developed by 711 HPW/RHA manage communications from the simulation suite to the FTU's AF-DCGS equipment. AFSERS components feed the AF-DCGS systems information from the simulated ISR platforms. The proof-of-concept system is enabling the FTU AF-DCGS workstations to function in the classroom the same way workstations function operationally. Ongoing 711 HPW/RHAS research for Air Force ISR personnel sponsored by the Information Operations and Special Programs Branch has been essential to enabling the right partners to come together for this collaborative effort. The proof-of-concept system was installed in May 2007 and ownership transferred to 17 Training Support Squadron. 711 HPW/RHA has continuously improved the proof-of-concept system over the last year and continues to gain valuable data to help pave the way to bring DMO training and rehearsal capabilities to Air Force ISR personnel and validate new training methodologies and techniques.

Mr. Geoffrey Barbier, 711 HPW/RHAS

RHA Investigates Latest Gaming Technologies for Military Simulation

The 711 HPW/RHA has initiated a Gaming Technology Research and Development project with the goal of evaluating the full training potential of technologies. Gaming technology exploits the latest in computer hardware, pushing the envelope of visual graphics, usability and connectivity, while offering rapid development capabilities at low cost to the end-user. The use of Gaming technology for interactive military training has been hindered by the fidelity of models used in the commercial game engines. This deficiency can be overcome by driving the game environment with external, high fidelity, validated models. Researchers are investigating what levels of fidelity and correlation can be reached and whether increasing the fidelity of the existing games can improve training value.

A commercial-off-the-shelf flight simulation program, utilizing a powerful but low-cost software development kit and leveraging support from an extensive development community, was successfully integrated with a C-based computer generated forces/electronic warfare environment to run validated high fidelity models. Software plug-ins developed for the flight simulator enabled it to communicate with the military's Distributed Interactive Simulation network protocol, show threat information on a cockpit RADAR Warning Receiver scope, and model Unmanned Aerial Vehicle flight and camera actions. Research will continue into database cor-

Screen shot of commercial-off-the-shelf flight simulator with a RWR scope. Photo by Lt Clint Kam.

relation, hardware performance enhancements, and training effectiveness of the gaming systems.

Lt Clinton J. Kam, 711 HPW/RHAE

Science and Technology Areas of Relevance for AETC Future Learning Systems

There are three major areas of research and development underway at the Mesa Research Site that have relevance to the Air Education and Training Command (AETC) Future Learning Systems (FLS) capabilities. They are Continuous Learning for Aiding and Training Decision Making, Computational Replicates, and Multi-Modal Immersion. Continuous Learning for Aiding and Training Decision Making is unique in that it is the only program in the Air Force Research Laboratory conducting research to develop better methods for Live, Virtual, and Constructive (LVC) training, aiding, and rehearsing for individuals and teams. The goal is a seamless learning enterprise that can pro-

vide learners with knowledge to effectively perform their jobs anytime and anywhere. This work will also provide the capability to track learning and performance for individuals and teams and to tailor learning events for targeted improvements in performance and effectiveness. Partner research programs of merit at Mesa are Computational Replicates and Multi-Modal Immersion. The goal of the Computational Replicates program is to create new cognitive science-based technology options for the Air Force, including: synthetic teammates for constructive blue force representations, pedagogical agents for adaptive training and rehearsal systems, and analysis tools for warfighter performance optimization.

The goals of the Multi-Modal Immersion program are to develop and validate human-centered tetherless immersive training and aiding environments providing multiple modes of stimuli, enabling interaction with distributed LVC participants, entities, objects and/or information. The capabilities developed in all three of these research programs align directly with the stated goals of the AETC FLS and will be validated across multiple mission domains and applications (e.g., air, C4ISR, cyber, space).

Dr. Winston Bennett, 711 HPW/RHAS

TARGETS OF OPPORTUNITY

Mesa Researchers Lead Training Solution Analysis Team for Unmanned Aerial Systems

711 HPW/RHA researchers, in collaboration with the 11th Reconnaissance Squadron (11 RS) and the Air Force Safety Center analyzed Predator class A, B, and C mishaps to identify problem areas that appeared to have potential training solutions. Results from early work were presented in an InterService/Industry Training, Simulation, and Education Conference in December 2007. The paper, entitled *Birds of Prey: Training solutions to Human Factors Problems* highlighted Predator mishap data indicating the dynamic nature of mishaps over time indicating evolving human factors issues of relevance. The leading, training-related mishap causes in recent mishaps were channelized attention, no training for tasks attempted, and decision making/

risk assessment. A panel of expert A-10, F-16, F-15, and MQ-1 pilots reviewed and validated these findings. They identified channelized attention, task prioritization, and course of action selected as problem areas in all of these platforms, and a prioritized list of interventions to address these problems was developed based on feasibility and probable benefits. Enhanced academic content and game-based, hands-on training emerged as leading candidates. Work to develop and evaluate candidate solutions is currently underway via a Small Business Innovation Research effort. Crew Training International and Anacapa Sciences are working to add and evaluate these exemplars in the 11 RS curriculum. Enhanced student performance tracking in several training events

was developed to support this evaluation.

Historically, the Air Force used experienced, rated pilots or navigators as Predator operators and is currently considering candidates with alternative backgrounds, including recent undergraduate pilot training graduates and officers who are not rated. The enhanced performance measurement capability that was developed to assess the impacts of mishap reduction training interventions is also being used to provide student performance data supporting an Air Force Chief of Staff initiative to assess the impacts of training candidates with varying experience backgrounds.

Dr. Robert Nullmeyer, 711 HPW/RHAS

BRIEFS AND DEBRIEFS

Live, Virtual, and Constructive Demonstrations planned at the Nellis Test and Training Range

Starting in late FY09, researchers from 711 HPW/RHA have proposed to team with United States Warfare Center, the 98th Range Wing, Boeing, and Cubic Defense Applications for an operational demonstration of Live, Virtual, and Constructive (LVC) operations at Nellis. This demonstration will pave the way for operational LVC by demonstrating secure LVC data from a live aircraft to be sent bidirectionally to an LVC node at Nellis, using off-the-shelf technology from Cubic and AFRL/RHA gateway software. The data from all 3 environments (Live, Virtual, and Constructive) will be captured and analyzed real time. Longer term proposed efforts involve scaling up the LVC ops capabilities and tools, automating data analyses conducted today by Range Training Officers, and saving thousands of hours of shot reconstruction time, while providing flights with key performance measurement data for every mission. The initial demonstration will include a software modified F-15E from Boeing St. Louis and aggressor aircraft from Nellis. The modified aircraft will display LVC data on the radar, the radar warning receiver, the data link display, and the advanced targeting pod simultaneously.

Ms. Kristen Barrera, 711 HPW/RHAS

Bringing LVC Ops into 5th Generation Aircraft

Working with Air Combat Command, F-22 training development engineers, and members of the F-35 Office Advisory Group, members of the 711 HPW/RHAS, along with Boeing and Lockheed Martin Advanced Combat Simulator (ACS) Group, will investigate alternative solutions to bring Live, Virtual, and Constructive technology into the 5th generation training environment. With the tremendous training challenges these aircraft face, it's hoped that the addition of LVC technology will provide better and more realistic training opportunities, precise performance measurement capabilities, proficiency and performance-based debriefing, and significant cost savings to the Combat Air Forces.

Ms. Kristen Barrera, 711 HPW/RHAS, Mr. Robert Rickard, 711 HPW/RHA

711 HPW/RHAS participates with Boeing on Project Alpine 2

In November, members of the 711 HPW/RHAS participated in a live flight demonstration of LVC Operations with software-modified F-15E from Boeing St. Louis. The aircraft simultaneously displayed LVC data on the radar, the radar warning receiver, the data link display, and the advanced targeting pod. The modified aircraft flew with an F-15E simulator on three operational flight profiles and demonstrated the tremendous training advancement opportunities that LVC provided in both 4th and 5th generation fighter aircraft. 711 HPW/RHA personnel had integrated the Division's recently completed CAT 1 Advanced Technology Demonstration performance evaluation and tracking technology with the Boeing system and recorded, analyzed, and provided debrief data both real time and post mission to the demonstration.

Ms. Kristen Barrera, 711 HPW/RHAS

Two New Hires Advancing Scientific Frontiers in Cognitive Models and Agents

The mission of the Cognitive Models and Agents Branch (711 HPW/RHAC) is to research, develop, and demonstrate leading edge technologies and innovative cognitive models that support the evolution of the global decision environment. 711 HPW/RHAC also administers the Night Vision Operations Center of Excellence. The branch's core in-house research effort is the creation of Computational Replicates, one of RHA's Focused Long-Term Challenge product lines. Along with Immersive Environments and Continuous Learning, Computational Replicates will enable the far-term vision for Live, Virtual and Constructive (LVC) operations. In this issue of Fight's On! we highlight two of RHAC's recent hires, Dr. Tiffany Jastrzembski and Dr. Scott Douglass, both of whom already are contributing at a high level to the scientific and technical foundation we need for the Computational Replicates.

Dr. Tiffany Jastrzembski was recognized this year by her peers in the scientific community with two distinguished awards for research conducted while she was a graduate student pursuing her Ph.D. in Cognitive Psychology at Florida State University. First, the American Psychological Association (APA), Division of Experimental Psychology, awarded Dr. Jastrzembski a New Investigator Award for an article in the *Journal of Experimental Psychology: Applied* stemming from her dissertation research. The article was published in 2007 and was titled, "The Model Human Processor and the Older Adult: Parameter Estimation and Validation within a Mobile Phone Task." This award recognizes her contributions to the fields of human factors engineering, cognitive modeling, and cognitive aging. Her dissertation demonstrated that age-sensitive processing parameters are valid for cognitive modeling purposes, can help designers understand age-related performance across different interface designs, and may support development of age-sensitive technologies. Second, Dr. Jastrzembski was honored with the 2008 Best Ergonomics in Design Article Award by the Human Factors and Ergonomics Society, for her article entitled "What Older Adults Can Teach Us About Designing Better Ballots." This research was funded as a student project through the multi-university Center for Research and Education on Aging and Technology Enhancement Program, as a side project during her doctoral work at Florida State. This award recognizes her contributions to the fields of human factors, cognitive aging, and voting design. Her research findings demonstrate that the application of a gerontechnological approach to voting design (i.e., designing with the older population in mind), can minimize errors and increase efficiency for users of all ages, which in turn helps minimize wait times at the polls and decreases the number of spoiled ballots. Congratulations to Dr. Jastrzembski for these multiple awards!

Dr. Scott Douglass joined Team Mesa last November after successfully defending his Ph.D. in Cognitive Psychology at Carnegie Mellon University. This spring and summer he worked with Dr. David Luginbuhl, who manages the Air Force Office of Scientic Research's (AFOSR) Software and Systems Program, to co-organize and co-chair a joint AFOSR-RHA workshop titled Cognitive Modeling and Software Engineering: Synergistic Approaches to Representing Human Behavior. During the two-day event, attendees from academia, industry, and various government agencies were briefed by 19 members of the software engineering and cognitive modeling communities. Workshop presentations and follow-up discussions explored the overlap between the methodologies and objectives of these two communities. The briefings and discussions indicated that cognitive modeling and software engineering are traveling down similar paths. Both are trying to develop explanations and simulations of radically complex systems. Both are also finding that their current specification and representation languages are inadequate for their respective modeling and system specification needs. While the impact and possible collaborative outcomes of the workshop are still being assessed, activities during the event succeeded in highlighting potential synergies between the two fields. Follow-up to the workshop will further explore: (1) how human-centered systems design might benefit from cognitive modeling; and (2) how cognitive modelers building large-scale models might benefit from software engineering. The workshop will hopefully act as a catalyst that fosters a fusion of assets through which the cognitive modeling and software engineering communities will learn from each other, combine expertise, and attack their shared problem. Synergies between the software engineering and cognitive modeling communities will hopefully facilitate progress in ongoing basic and applied research efforts supporting AFRL's long-term technology goals.

Dr. Kevin Gluck, 711 HPW/RHAC

CPAS retrieves targeting information from JADOCS and chat information from IRC linking the data together on a time line to reconstruct the command and control process of dynamic effects planning. Research data collected during T-REX 09-1 will guide potential modifications to chat TTPs prior to publication. RHA scientists will also continue to develop and mature CIFTS and CPAS into products for transitions to trainers to help assess individual and team performance.

T-REX continues to provide a forum to test new training methodologies and technologies that will enhance warfighter training, making them better prepared to fight today's war. Data collected through assessment systems and warfighter feedback will provide RHA scientists with valuable insight in analysis of team performance at the operational level of warfare and transition effective training methods to Air Combat Command for incorporation into AOC training worldwide.

Lt Andrea Wolfe, 711 HPW/RHAS
Mr. Oscar Garcia, 711 HPW/RHAS
Mr. Todd Denning, 711 HPW/RHA